

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 06:16:43 ; Search time 80 Seconds  
(without alignments)  
8303.453 Million cell updates/sec

Title: US-10-063-730-115  
Perfect score: 1197  
Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaa 1197

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 10

Total number of hits satisfying chosen parameters: 257976

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6C COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1057	88.3	1380	4	US-09-620-312D-387
2	547	45.7	1428	4	US-09-620-312D-386
3	26	2.2	11050	4	US-10-204-708-86
4	25	2.1	87350	3	US-08-781-891-79
5	25	2.1	87350	4	US-09-618-166-79
6	25	2.1	87543	4	US-09-791-211-3
7	25	2.1	162450	4	US-09-345-882-1
8	24	2.0	414	1	US-08-377-687-48
9	24	2.0	414	1	US-08-777-192-48
10	24	2.0	414	3	US-08-971-982-48
11	24	2.0	414	4	US-09-077-951-19
12	24	2.0	414	4	US-09-077-948A-45
13	24	2.0	705	3	US-09-328-111-674
14	24	2.0	1396	1	US-08-123-161A-11
15	24	2.0	1396	1	US-08-483-278-11
16	24	2.0	1529	3	US-09-189-760-5
17	24	2.0	1529	3	US-09-188-811-5
18	24	2.0	1529	3	US-09-514-422-5
19	24	2.0	2017	1	US-07-667-276A-3
20	24	2.0	2041	4	US-09-149-476-131
21	24	2.0	2072	4	US-09-833-381-1791
22	24	2.0	2328	4	US-08-811-481-34
23	24	2.0	2328	4	US-09-876-527-34
24	24	2.0	2494	3	US-09-189-760-1
25	24	2.0	2494	3	US-09-514-422-1
26	24	2.0	13104	3	US-08-256-799-4
27	24	2.0	13104	3	US-08-462-437-4

28	23	1.9	55	2	US-08-771-624B-8	Sequence 8, Appli
29	23	1.9	104	4	US-09-621-976-9392	Sequence 9392, Ap
30	23	1.9	553	4	US-09-148-545-133	Sequence 133, App
31	23	1.9	557	4	US-09-148-545-80	Sequence 80, Appl
32	23	1.9	653	4	US-09-373-750-1	Sequence 1, Appli
c 33	23	1.9	976	1	US-08-125-628-7	Sequence 7, Appli
c 34	23	1.9	976	1	US-08-125-628-8	Sequence 8, Appli
35	23	1.9	990	4	US-09-461-325-94	Sequence 94, Appl
36	23	1.9	990	4	US-10-012-542-94	Sequence 94, Appl
37	23	1.9	1027	4	US-09-593-887-9	Sequence 9, Appli
38	23	1.9	1296	4	US-09-533-029-63	Sequence 63, Appl
39	23	1.9	1348	4	US-09-152-060-14	Sequence 14, Appl
40	23	1.9	1500	3	US-09-052-778-1	Sequence 1, Appli
41	23	1.9	1771	4	US-09-533-029-103	Sequence 103, App
42	23	1.9	1778	4	US-09-833-381-2044	Sequence 2044, Ap
43	23	1.9	2114	4	US-09-130-491-7	Sequence 7, Appli
c 44	23	1.9	2285	4	US-09-230-225B-1	Sequence 1, Appli
45	23	1.9	17949	3	US-09-087-465-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-620-312D-387  
; Sequence 387, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 387  
; LENGTH: 1380  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (131)..(1084)  
US-09-620-312D-387

Query Match 88.3%; Score 1057; DB 4; Length 1380;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1107; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 73 TCAGAGAAATGTGAAGACTGTACATCTTAAATGCAGAGCTTTTAAATCCAGAAAAAT 132  
Db 145 TCAGAGAAATGTGAAGACTGTACATCTTAAATGCAGAGCTTTTAAATCCAGAAAAAT 204

```
QY 133 ATGTAATCACTTAAGATTGTGGACTGGTGTGTTTGGTATCCTGGCCCTAACTCTAATTGT 192
Db 205 ATGTAATCACTTAAGATTGTGGACTGGTGTGTTTGGTATCCTGGCCCTAACTCTAATTGT 264
QY 193 CTTGTTTGGGGAGCAAGCACTTCTGGCGGGAGGTACCCAAAAGCCATGACATGGA 252
Db 265 CTTGTTTGGGGAGCAAGCACTTCTGGCGGGAGGTACCCAAAAGCCATGACATGGA 324
QY 253 GCACACTTTCTACAGCAATGGAGAGAGAGATTTCATGGAATTGATCCTCTGAC 312
Db 325 GCACACTTTCTACAGCAATGGAGAGAGAGATTTCATGGAATTGATCCTCTGAC 384
QY 313 CAGAACTGAAATATTCAGAAAGCGGAATGSCACTGATGAAACATTTGGAAGTGACGACTT 372
Db 385 CAGAACTGAAATATTCAGAAAGCGGAATGSCACTGATGAAACATTTGGAAGTGACGACTT 444
QY 373 TAAAAACGATACACTGGGATCTACTTCTGGGTCTTCAAAAATGTTTTATCAAACTCA 432
Db 445 TAAAAACGATACACTGGGATCTACTTCTGGGTCTTCAAAAATGTTTTATCAAACTCA 504
QY 433 GATTAAGTGATTCTCTGAATTTTCTGAAACAGAAAGGAAATAGATGAGAAATGAAGAAAT 492
Db 505 GATTAAGTGATTCTCTGAATTTTCTGAAACAGAAAGGAAATAGATGAGAAATGAAGAAAT 564
QY 493 TACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCAGCAGCAAAAAGCCTATTGAAA 552
Db 565 TACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCAGCAGCAAAAAGCCTATTGAAA 624
QY 553 CCGAGATTTTCTTAAAAATCCAAAATCTCGAGATTTGTGATAAGCGTGACCATGTATTG 612
Db 625 CCGAGATTTTCTTAAAAATCCAAAATCTCGAGATTTGTGATAAGCGTGACCATGTATTG 684
QY 613 GATCAATCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGGAAGA 672
Db 685 GATCAATCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGGAAGA 744
QY 673 TCTTCACTTTCTTCGCAACGAAAAAAGGATTTGAACAAAATGAACAGTGGGTGTGCC 732
Db 745 TCTTCACTTTCTTCGCAACGAAAAAAGGATTTGAACAAAATGAACAGTGGGTGTGCC 804
QY 733 TCAAGTGAAGTAGAAGAACCGGTACGCCAGCAAGCAAGTGAAGAGAACTTCCAAT 792
Db 805 TCAAGTGAAGTAGAAGAACCGGTACGCCAGCAAGCAAGTGAAGAGAACTTCCAAT 864
QY 793 AAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATCAGAGAGTTATTG 852
Db 865 AAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATCAGAGAGTTATTG 924
QY 853 TTGTATTTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTAGG 912
Db 925 TTGTATTTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTAGG 984
QY 913 CTACTACCCATATCCATCTACTTACCAGGAGACGAGTCATCTGCTGTCATCATGCC 972
Db 985 CTACTACCCATATCCATCTACTTACCAGGAGACGAGTCATCTGCTGTCATCATGCC 1044
QY 973 TTGTAACTGGTGGTGGCCCGATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAAT 1032
Db 1045 TTGTAACTGGTGGTGGCCCGATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAAT 1104
QY 1033 GCTTAACTGCTGGCAACATATAATAATGCATGCTATTCAATGAATTTCTGCTATGAG 1092
Db 1105 GCTTAACTGCTGGCAACATATAATAATGCATGCTATTCAATGAATTTCTGCTATGAG 1164
QY 1093 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATTACTTTGAGTAAATTCCTCTCTTCATG 1152
Db 1165 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATTACTTTGAGTAAATTCCTCTCTTCATG 1224
QY 1153 TTCTAATAAACTTCTATCAATTATCAACAA 1180
Db 1225 TTCTAATAAACTTCTATCAATTATCAACAA 1252
```

```
RESULT 2
US-09-620-312D-386
; Sequence 386, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 386
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)...(1132)
US-09-620-312D-386

Query Match 45.7%; Score 547; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 8.8e-240;
Matches 547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 634 AGTTCTGAGTTACAGACTTTGAGGAGGAGGAGAGATCTTCACCTTTCCTGCCAACGA 693
Db 754 AGTTCTGAGTTACAGACTTTGAGGAGGAGGAGAGATCTTCACCTTTCCTGCCAACGA 813
QY 694 AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAC 753
Db 814 AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAC 873
QY 754 CCGTACGCCAGCAAGCAAGTAGAGAGAACTTCCAAATAAATGACTATCTGAAATGG 813
Db 874 CCGTACGCCAGCAAGCAAGTAGAGAGAACTTCCAAATAAATGACTATCTGAAATGG 933
QY 814 AATAGAAATTTGATCCCATCTGATCGATGAGAGAGTTATTGTGTATTACTGCCGTCGAGG 873
Db 934 AATAGAAATTTGATCCCATCTGATCGATGAGAGAGTTATTGTGTATTACTGCCGTCGAGG 993
QY 874 CAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTG 933
Db 994 CAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTG 1053
QY 934 CTACCAAGGAGACGAGTCATCTGTCGTCATCATGCTTGTAACTGTTGTTGGTGGCCCG 993
Db 1054 CTACCAAGGAGACGAGTCATCTGTCGTCATCATGCTTGTAACTGTTGTTGGTGGCCCG 1113
QY 994 CATGCTGGGAGGGTCTAATAGAGGTTTGAAGCTCAAAATGCTTAAATGCTGCTGCAACATA 1053
Db 1114 CATGCTGGGAGGGTCTAATAGAGGTTTGAAGCTCAAAATGCTTAAATGCTGCTGCAACATA 1173
```

QY 1054 TAATAATGATGCTATTCAATGAATTCGCTATGAGGCATCGCCCGCTGGTAGCCA 1113  
Db 1174 TAATAATGATGCTATTCAATGAATTCGCTATGAGGCATCGCCCGCTGGTAGCCA 1233  
QY 1114 GCCTCCAGAACTACTGTAGTAGTAATCCCTCTCTCATGTTCTTAATAAACTTCTACATTA 1173  
Db 1234 GCCTCCAGAACTACTGTAGTAGTAATCCCTCTCTCATGTTCTTAATAAACTTCTACATTA 1293  
QY 1174 TCACCAA 1180  
Db 1294 TCACCAA 1300

RESULT 3  
US-10-204-708-86/c  
; Sequence 86, Application US/10204708  
; Patent No. 6677731  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; TITLE OF INVENTION: by Assessing DNA Methylation  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/204,708  
; CURRENT FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 86  
; LENGTH: 11050  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-204-708-86

Query Match 2.2%; Score 26; DB 4; Length 11050;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1171 TTATCACCACCAAAAAAAAAAAAAAAAAAAAA 1196  
Db 744 TTATCACCACCAAAAAAAAAAAAAAAAAAAAA 719

RESULT 4  
US-08-781-891-79/c  
; Sequence 79, Application US/08781891  
; Patent No. 6090620  
; GENERAL INFORMATION:  
; APPLICANT: Fu, Ying-Hui  
; APPLICANT: Yu, Chang-En  
; APPLICANT: Oshima, Junko  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Schellenberg, Gerald D.  
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
; TITLE OF INVENTION: WERNER'S SYNDROME  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington

COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,891  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6090620tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 240052.419  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-781-891-79  
Query Match 2.1%; Score 25; DB 3; Length 87350;  
Best Local Similarity 100.0%; Pred. No. 0.048;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1172 TATCACCACCAAAAAAAAAAAAAAAAAAAAA 1196  
Db 29899 TATCACCACCAAAAAAAAAAAAAAAAAAAAA 29875  
RESULT 5  
US-09-618-166-79/c  
; Sequence 79, Application US/09618166  
; Patent No. 6583112  
; GENERAL INFORMATION:  
; APPLICANT: Fu, Ying-Hui  
; APPLICANT: Yu, Chang-En  
; APPLICANT: Oshima, Junko  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Schellenberg, Gerald D.  
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
; TITLE OF INVENTION: WERNER'S SYNDROME  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Seed Intellectual Property Law Group  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/618,166  
FILING DATE: 17-Jul-2000  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 240052.419C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:



[illegible]

```
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
```

```
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match          2.1%; Score 25; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1173 ATCACCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 151927 ATCACCACCAAAAAAAAAAAAAAAAAAAAA 151903

RESULT 8
US-08-377-687-48
; Sequence 48, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSEBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```

;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
US-08-377-687-48

Query Match      2.0%; Score 24; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 9
US-08-777-192-48
; Sequence 48, Application US/08777192
; Patent No. 5824869
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-08-777-192-48

Query Match      2.0%; Score 24; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 11
US-09-077-951-19
; Sequence 19, Application US/09077951
; Patent No. 6372888
; GENERAL INFORMATION:
; APPLICANT: De Samblaux, Genoveva
; APPLICANT: Broekaert, Willem
; APPLICANT: Rees, Sarah
; TITLE OF INVENTION: Antifungal Proteins
; FILE REFERENCE: PPD50093
; CURRENT APPLICATION NUMBER: US/09/077,951

```

```
; CURRENT FILING DATE: 1999-03-11
; EARLIER APPLICATION NUMBER: GB 9525474.4
; EARLIER FILING DATE: 1995-12-13
; EARLIER APPLICATION NUMBER: PCT/GB96/03065
; EARLIER FILING DATE: 1996-12-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Raphanus sativus
US-09-077-951-19

Query Match
Best Local Similarity 2.0%; Score 24; DB 4; Length 414;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 12
US-09-077-948A-45
; Sequence 45, Application US/09077948A
; Patent No. 6605698
; GENERAL INFORMATION:
; APPLICANT: Van Amerongen, Aart
; APPLICANT: Fant, Franky
; APPLICANT: Borremans, Frans
; APPLICANT: De Samblaux, Genevieve
; APPLICANT: Sitjtsma, Lolke
; APPLICANT: Meloen, Robbert
; APPLICANT: Puijk, Wouter
; APPLICANT: Schaeper, Wilhelmus
; APPLICANT: Broekaert, Willem
; APPLICANT: Van Gelder, Wilhelmus
; APPLICANT: Rees, Sarah
; TITLE OF INVENTION: Antifungal Proteins
; FILE REFERENCE: 109846-257(SYN-035)
; CURRENT APPLICATION NUMBER: US/09/077,948A
; CURRENT FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/GB96/03068
; PRIOR FILING DATE: 1996-12-12
; PRIOR APPLICATION NUMBER: GB 9606552.9
; PRIOR FILING DATE: 1996-03-28
; PRIOR APPLICATION NUMBER: GB 9525455.3
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Raphanus sativus
US-09-077-948A-45

Query Match
Best Local Similarity 2.0%; Score 24; DB 4; Length 414;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 13
US-09-328-111-674/c
; Sequence 674, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Ascle, Jon H.
```

```
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 674
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(705)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-674

Query Match
Best Local Similarity 2.0%; Score 24; DB 3; Length 705;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 35 TCACCAAAAAAAAAAAAAAAAAAAAA 12

RESULT 14
US-08-123-161A-11
; Sequence 11, Application US/08123161A
; Patent No. 5449616
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Roberts, Steven L.
; APPLICANT: Anderson, Richard D.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Yang, Bin
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,161A
; FILING DATE: 16-SEP-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/946,234
; FILING DATE: 14-SEP-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIRF89-11A4
```



## TELECOMMUNICATION INFORMATION:

TELEPHONE: (207) 363-0558

TELEFAX: (207) 363-0528

INFORMATION FOR SEQ ID NO: 11:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1396 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 4..1164

US-08-123-161A-11

Query Match

Best Local Similarity 2.0%; Score 24; DB 1; Length 1396;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197

|||||

Db 1366 TCACCAAAAAAAAAAAAAAAAAAAAA 1389

## RESULT 15

US-08-483-278-11

; Sequence 11, Application US/08483278

; Patent No. 5686073

## GENERAL INFORMATION:

; APPLICANT: Campbell, Kevin P.

; APPLICANT: Ibraghimov, Oxana B.

; APPLICANT: Ervasti, James M.

; APPLICANT: Leveille, Cynthia J.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kevin M. Farrell, P.C.

; STREET: P.O. Box 999

; CITY: York Harbor

; STATE: ME

; COUNTRY: USA

; ZIP: 03911

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,278

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/123,161

; FILING DATE: 16-SEP-93

; ATTORNEY/AGENT INFORMATION:

; NAME: Farrell, Kevin M.

; REGISTRATION NUMBER: 35,505

; REFERENCE/DOCKET NUMBER: UIRF89-11A5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (207) 363-0558

; TELEFAX: (207) 363-0528

; INFORMATION FOR SEQ ID NO: 11:

## SEQUENCE CHARACTERISTICS:

; LENGTH: 1396 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 4..1164

; US-08-483-278-11

Query Match 2.0%; Score 24; DB 1; Length 1396;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197

|||||

Db 1366 TCACCAAAAAAAAAAAAAAAAAAAAA 1389

Search completed: September 3, 2004, 07:45:39

Job time : 82 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 04:36:32 ; Search time 80 seconds  
(without alignments)  
8303.453 Million cell updates/sec

Title: US-10-063-730-115

Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1174	98.1	1380	4	US-09-620-312D-387
2	1116	93.2	1428	4	US-09-620-312D-386
3	1111	9.3	1006	1	US-08-241-465B-3
4	109.4	9.1	1006	1	US-08-241-465B-2
5	106.2	8.9	1006	1	US-08-241-465B-6
6	104.6	8.7	1006	1	US-08-241-465B-5
7	90.8	7.6	892	1	US-08-241-465B-4
8	90	7.5	363	1	US-08-047-033-6
9	90	7.5	364	1	US-08-047-033-5
10	90	7.5	365	1	US-08-047-033-4
11	90	7.5	366	1	US-08-047-033-3
12	86	7.2	892	1	US-08-241-465B-7
13	53.2	4.4	7218	1	US-08-232-463-14
14	49.2	4.1	230	1	US-08-047-033-7
15	41.9	3.5	12886	4	US-09-453-702B-14
16	41	3.4	505	4	US-09-621-976-15639
17	39	3.3	1664976	4	US-08-916-421B-1
18	36.4	3.0	323	4	US-09-621-976-10374
19	36	3.0	1485	4	US-09-573-906-1
20	35.8	3.0	246	4	US-09-134-001C-2521
21	35.6	3.0	1506	4	US-09-134-001C-1278
22	35.2	2.9	29485	4	US-09-785-381-6
23	35	2.9	5183	1	US-08-459-568-3
24	35	2.9	5183	2	US-08-399-411-3
25	35	2.9	5868	3	US-08-516-859A-3
26	35	2.9	5868	4	US-09-586-472-3
27	35	2.9	5868	4	US-09-528-706-3

c	28	34.8	2.9	5562	4	US-10-204-708-64	Sequence 64, Appl
	29	34.8	2.9	1664976	4	US-08-916-421B-1	Sequence 1, Appl
	30	34.6	2.9	2241	4	US-09-023-942A-9	Sequence 9, Appl
	31	34.4	2.9	1176	3	US-09-072-384-16	Sequence 16, Appl
	32	34.2	2.9	289	3	US-09-007-005-17	Sequence 17, Appl
	33	34.2	2.9	289	3	US-09-244-796-17	Sequence 17, Appl
	34	34	2.8	1335	4	US-09-543-681A-1183	Sequence 1183, Ap
	35	34	2.8	1686	4	US-09-134-001C-1354	Sequence 1354, Ap
	36	34	2.8	2065	3	US-09-370-473-5	Sequence 5, Appl
	37	33.8	2.8	832	4	US-09-134-000C-45	Sequence 45, Appl
c	38	33.8	2.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
	39	33.8	2.8	1754	4	US-09-604-978-1	Sequence 1, Appl
	40	33.8	2.8	1754	4	US-09-604-728-1	Sequence 1, Appl
	41	33.6	2.8	1611	6	5213972-6	Patent No. 5213972
	42	33.6	2.8	3300	4	US-09-268-347-31	Sequence 31, Appl
c	43	33.6	2.8	8607	4	US-10-204-708-72	Sequence 72, Appl
	44	33.4	2.8	148	4	US-09-621-976-8100	Sequence 8100, Ap
	45	33.4	2.8	565	4	US-09-589-733C-6	Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-09-620-312D-387  
; Sequence 387, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Dmanac, Radje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 387  
; LENGTH: 1380  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (131)..(1084)  
US-09-620-312D-387

Query Match 98.1%; Score 1174; DB 4; Length 1380;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1177; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 CAGCAGTGTCCTCAGTCCTCTCAAAGCAAGAGTACTGTGCTGAGACCAT 60  
DB 73 CAGCAGGGGTCTCTCAGTCCTCTCAAAGCAAGAGTACTGTGCTGAGACCAT 132



QY 361 AGTCAGGACTTTAAAAACGGATACACTGGCATCTACTTCCTGGGTCTTCAAAATGTTT 420  
Db |||||  
QY 433 AGTACACGACTTTAAAAACGGATACACTGGCATCTACTTCCTGGGTCTTCAAAATGTTT 492  
Db |||||  
QY 421 TATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACAGAGGAGAAATAGATGA 480  
Db |||||  
QY 493 TATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACAGAGGAGAAATAGATGA 552  
QY 481 GAATGAAGAAATACCACTTCTTTTGACAGTCACTGATTTGGTCCCGACAGAAA 540  
Db |||||  
QY 553 GAATGAAGAAATACCACTTCTTTTGACAGTCACTGATTTGGTCCCGACAGAAA 612  
QY 541 GCCTATTGAAACCGAGATTTCTTTAAAAATTCAAAAATTCGAGAGATTTGTGATAAGT 600  
Db |||||  
QY 613 GCCTATTGAAACCGAGATTTCTTTAAAAATTCAAAAATTCGAGAGATTTGTGATAAGT 672  
QY 601 GACCATGTNTGGATCAATCCACTCTAAATNC----- 633  
Db |||||  
QY 673 GACCATGTNTGGATCAATCCACTCTAAATCAATATCAGGAACATTTGCAAGCAGTTGCATCA 732  
QY 634 -----AGTTTCTGAGTTACAGACTTTTCAGGAGGAGGAGGA 672  
Db |||||  
QY 733 CAACITTTGCATTTATCTTAGTTTCTGAGTTTACAGACTTTGAGGAGGAGGAGGA 792  
Db |||||  
QY 673 TCTTCACTTTCTGCCCAACGAAAAAAGGATTTGAACAAAAAGCAAGTGGTGGTCCC 732  
Db |||||  
QY 793 TCTTCACTTTCTGCCCAACGAAAAAAGGATTTGAACAAAAAGCAAGTGGTGGTCCC 852  
QY 733 TCAAGTGAAGTAGAGAGACCGGTACGCCAGACAAAGCAAGTGAAGAACTTCCAAT 792  
Db |||||  
QY 853 TCAAGTGAAGTAGAGAGACCGGTACGCCAGACAAAGCAAGTGAAGAACTTCCAAT 912  
QY 793 AAATGACTACTGAAATGGAATAGAAATTTGATCCCATGCTGATCAGAGAGTTATTG 852  
Db |||||  
QY 913 AAATGACTACTGAAATGGAATAGAAATTTGATCCCATGCTGATCAGAGAGTTATTG 972  
QY 853 TTGTATTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTTAGG 912  
Db |||||  
QY 973 TTGTATTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTTAGG 1032  
QY 913 CTACTACCATATCCATCTACTGCTACCAAGGAGGAGCAGTCACTGCTGTCATCATGCC 1092  
Db |||||  
QY 1033 CTACTACCATATCCATCTACTGCTACCAAGGAGGAGCAGTCACTGCTGTCATCATGCC 1092  
QY 973 TTGTAACTGTTGGTGGCGCGCATGCTGGGAGGCTCTAATAGGAGTTTGGAGTCAAT 1032  
Db |||||  
QY 1093 TTGTAACTGTTGGTGGCGCGCATGCTGGGAGGCTCTAATAGGAGTTTGGAGTCAAT 1152  
QY 1033 GCTTAACTGCTGGCAACATATAATAATGCATGCTATTCAATGAATTTCTGCTATGAG 1092  
Db |||||  
QY 1153 GCTTAACTGCTGGCAACATATAATAATGCATGCTATTCAATGAATTTCTGCTATGAG 1212  
QY 1093 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATCTTGTAGTAAATCTCTCTTCATG 1152  
Db |||||  
QY 1213 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATCTTGTAGTAAATCTCTCTTCATG 1272  
QY 1153 TTCTAATAAACTTCTACATTTATCACCAGAA 1182  
Db |||||  
QY 1273 TTCTAATAAACTTCTACATTTATCACCAGAA 1302

RESULT 3  
US-08-241-465B-3  
; Sequence 3, Application US/08241465B  
; Patent No. 5719125  
; GENERAL INFORMATION:  
; APPLICANT: Fujio SUZUKI  
; APPLICANT: Yuji HIRAKI  
; APPLICANT: Kazuhiro TAKAHASHI  
; APPLICANT: Junko SUZUKI  
; APPLICANT: Jun KONDO  
; APPLICANT: Atsuko KOHARA

; APPLICANT: Akiko MORI  
; APPLICANT: Ei YAMADA  
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; COUNTRY: D.C.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,465B  
; FILING DATE: May 11, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1006 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:  
; ORGANISM: Human being (Homo Sapiens)  
; IMMEDIATE SOURCE:  
; CLONE: phCHM-16-3  
; FEATURE:  
; NAME/KEY: P CDS  
; LOCATION: 2..1003  
; IDENTIFICATION METHOD: E  
; US-08-241-465B-3

Query Match 9.3%; Score 111; DB 1; Length 1006;  
Best Local Similarity 49.8%; Pred. No. 3.7e-32;  
Matches 412; Conservative 0; Mismatches 395; Indels 24; Gaps 4;  
QY 203 GGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAAGCCCTATGACATGGAGCACACTTTC 262  
Db |||||  
QY 179 GGGGCTTCTACTTCTGGAAGGGGAGCGAGTCACATTTACAATGTCCTATACACCATG 238  
Db |||||  
QY 263 TACAGCAATGGAGAGAGAAAGATTTCATGGAATTCCTGTCGACGAACTGAA 322  
Db |||||  
QY 239 AGTATCAATGGGAACTTACAAATGGGTCAATGGAATAGACGCTGGGAACAACCTTGGAG 298  
QY 323 ATATTACAGAGCGAATGCACTGATGAACATTCGAAGTGCACGACTTAAAAACGGA 382  
Db |||||  
QY 299 ACCTTTAAATGGGAATGGAGCTGAAGAGCAATTCAGTTTAAATGATTTCCAGATGGC 358  
QY 383 TACATGGCATCTACTTCGTGGGTCTTCAAAATGTTTATCAAACTCAGATTTAAAGTG 442  
Db |||||  
QY 359 ATCAGAGGAATTCGTTTGTCTGGAGGAGAGAGTGTACATTTAAGGCGAGTGAAGGCT 418  
QY 443 ATT-----CCTGAATTTCTGAACAGAGGAAATAGATGAATGAAGAA----- 490  
Db |||||  
QY 419 CGTATTCCTGAGTGGCGCGGTGACCAACAGAGCATCTCTCCAAACTGGGAAGCAAG 478  
QY 491 ---ATTACCACTTCTTTGACAGTCACTGATTTGGTCCCGACAGAAAGCCTATT 547  
Db |||||  
QY 479 ATCATCCAGTCAATATGAAGAAATTTCTTATCTGGTGGCTGTAGATCAGCCTGTG 538  
QY 548 GAAACCGAGATTTCTTAAAAATTCAAAAATTCCTGAGATTTGTGATACGTCACCATG 607

Db 539 AAGGACAAAGCTTCTTG---AATTCTAAGGTGTAGAACTCTGGGTGACCTTCTTATT 595  
QY 608 TATTGGATCAATCCCACTTAATATCACTTTCTGAGTTTACAAAGCTTTGAGGAGGAGGA 667  
Db 596 TTCTGGCTTAACCACTTATCCAAAGAAATCCAGAGGAAAGAGAGAGTGTGAAGA 655  
QY 668 GAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 727  
Db 656 AAAATTGTTCCAACTACCACAAAAAGACACACAGTGGACCAAGAGCAACCCAGCGCT 715  
QY 728 GTCCCTCAAGTGAAGTAGAAGACCGT-----CAGCCAGACAAGCAAGTAGAGAA 781  
Db 716 GGAAGACTGAATAATGAACACCAAGACCCAGTGTTCAGAGGACTCAAGCCCTTCAATCCT 775  
QY 782 GAACCTTCAATAAATGACTATATCTGAAATGAATAGAAATTTGATCCCATGCTGATGAG 841  
Db 776 GATAATCTTATCATCAGAGAAAGGAAAGCATGACATTCGACCTTAGACTGGATCAC 835  
QY 842 AGAGTTATTGTTGATTACTTTCGCGTCGAGCAACCGCTATTGCGCGCGGTCTGTGAA 901  
Db 836 GAAGGAATCTGTTGTATAGATGTAGCGGAGCTACACCCACTGCCAGAAAGATCTGTGAA 895  
QY 902 CCTTTACTAGGCTACTACCATATCCATCTCTACCAAGGAGAGTCACTCTGCTGCT 961  
Db 896 CCCCTGGGGGGCTATTACCCATGGCCTTATAATATCAAGCTGCGCTTTCGGCGTGCAGA 955  
QY 962 GTCATCATGCTTGTAACTGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012  
Db 956 GTCATCATGCTATGATGTTGGTGGCGCGCATGCTGGGATGTTGTA 1006

## RESULT 4

US-08-241-465B-2  
; Sequence 2, Application US/08241465B  
; Patent No. 5719125  
; GENERAL INFORMATION:  
; APPLICANT: Fujio SUZUKI  
; APPLICANT: Yuji HIRAKI  
; APPLICANT: Kazuhiro TAKAHASHI  
; APPLICANT: Junko SUZUKI  
; APPLICANT: Jun KONDO  
; APPLICANT: Atsuko KOHARA  
; APPLICANT: Akiko MORI  
; APPLICANT: Ei YAMADA  
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-1 PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; COUNTRY: D.C.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,465B  
; FILING DATE: May 11, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1006 base pairs  
; TYPE: nucleic acid

; STRANDEDNESS: double  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:  
; ORGANISM: Human being (Homo Sapiens)  
; IMMEDIATE SOURCE:  
; CLONE: phCM-13-6  
; FEATURE:  
; NAME/KEY: P CDS  
; LOCATION: 2..1003  
; IDENTIFICATION METHOD: E  
US-08-241-465B-2  
  
Query Match 9.1%; Score 109.4; DB 1; Length 1006;  
Best Local Similarity 49.5%; Pred. No. 1.1e-21;  
Matches 411; Conservative 0; Mismatches 396; Indels 24; Gaps 4;  
  
QY 203 GGGAGCAGCACTTCTGGCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTTTC 262  
Db 179 GGGGCTTCTACTTCTGGAAGGGAGCGACAGTCACATTTACATGTCCATTACACCATG 238  
QY 263 TACGCAATGGAGAGAAGAAGATTTCATGGAAATTTGATCCTGTGACCAAGAACTGAA 322  
Db 239 AGTATCAATGGGAAACTACAGATGGGTCAATGGAAATAGACGCTGGGAAACAACCTGGAG 298  
QY 323 ATATTCAGAGCGGAAATGGCACTCATGAACATTTGGAAGTGCACGACTTTAAAAACGGA 382  
Db 299 ACCTTTAAATGGGAAGTGGAGCTGAAGAAGCAATTGCAGTTTAATGATTTCCAGATGCG 358  
QY 383 TACACTGGCATCTACTTCTGGGTCTTCAAAAAATGTTTATCAAAACTCAGATTAAGTG 442  
Db 359 ATCAGAAATTCGTTTGTCTGGAGGAGAGTCTACATTTAAAGCGCAAGTGAAGCT 418  
QY 443 ATT-----CCTGAATTTTCTGAACAGAGAGAGAAATAGATGAGAAATGAAGAA---- 490  
Db 419 CGTATTCCTGAGGTGGCGCGCTGACCAACAGAGCATCTCTCCAAACTGGAAGCAAG 478  
QY 491 ---ATTACCAACACTTCTTTGAAACAGTCAGTGAATTTGGTCCCGACAGAAAGCCTATT 547  
Db 479 ATCATGCCAGTCAAAATATGAAGAAAATTTCTTATCTGGTGGCTGTAGATCAGCCTGTG 538  
QY 548 GAAACCGGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATACCTGACCATG 607  
Db 539 AAGGACAGAGCTTCTTGAG---TTCTAAGGTGTTAGAACTCTGCGGTGACCTTCTTATT 595  
QY 608 TATTGGATCAATCCCACTTAATATCAGTTTCTGAGTTTACAAGACTTTGAGAGGAGGA 667  
Db 596 TTCTGGCTTAACCAACCTTATCCAAAGAAATCCAGAGGAAAGAGAGAGTGTGAAGA 655  
QY 668 GAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 727  
Db 656 AAAATTGTTCCAACTACCACAAAAAGACACACAGTGGACCGAGCAACCCAGCGCT 715  
QY 728 GTCCCTCAAGTGAAGTAGAAGACCGCT-----CAGCCAGACAAGCAAGTAGAGAA 781  
Db 716 GGAAGACTGAATAATGAACACCAAGCCAGTGTTCAGAGGACTCACAAGCCTTCAATCCT 775  
QY 782 GAACCTTCAATAAATGACTATATCTGAAATGGAATGAAATTTGATCCCATGCTGATGAG 841  
Db 776 GATAATCTTATCATCAGCAGGAAAGGGAAGCATGACATTCGACCTTAGACTGGATCAC 835  
QY 842 AGAGGTATTGTTGATTACTTTCGCGTCGAGGCAACCGCTATTGCGCGCGCTCTGTGAA 901  
Db 836 GAAGGAATCTGTTGTATAGAAATGTAGCGGAGCTACACCATGTCAGAGAAATCTGTGAA 895  
QY 902 CCTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGGAGGAGTCACTCTCTCGT 961  
Db 896 CCCCTGGGGGGCTATTACCCATGGCCTTATAATATCAAGGTGCGCTTTCGGCGTGCAGA 955  
QY 962 GTCATCATGCTTGTAACTGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012  
Db 956 GTCATCATGCTATGATGTTGGTGGCGCGCATCTCTGGGATGTTGTA 1006

```
RESULT 5
US-08-241-465B-6
; Sequence 6, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI
; APPLICANT: YUJI HIRAKI
; APPLICANT: KAZUHIRO TAKAHASHI
; APPLICANT: JUNKO SUZUKI
; APPLICANT: JUN KONDO
; APPLICANT: ATSUKO KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,465B
; FILING DATE: May 11, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
US-08-241-465B-6

Query Match      8.9%; Score 106.2; DB 1; Length 1006;
Best Local Similarity 49.2%; Pred. No. 9e-21;
Matches 409; Conservative 0; Mismatches 398; Indels 24; Gaps 4;

QY 203 GGGAGCAAGACATCTTGGCCGGAGGTACCCAAAAGCCCTATGACATGGAGCACACTTTC 262
    |||
DB 179 GGGGCCCTTCTACTCTTGGAGGGGAGGAGGACATGACATTTACATGTCCTTACACCATG 238
    |||

QY 263 TACAGCAATGGAGAGAGAGAGATTACATGGAATTTGATCCTGTGACCAAGAACTGAA 322
    |||
DB 239 AGTATCAATGGGAACTTACAGATGGGTCAATGGAAATAGACGCTGGGAACAACCTGGAG 298
    |||

QY 323 ATATTGAGAGCGGAATGACATGATGAACAATTGGAAGTGCACGACTTTAAAAACGGA 362
    |||
DB 299 ACCTTTAAAAATGGGAAGTGGAGCTGAAGAGCAATTGCAGTTAATGATTTCCAGAATGGC 358
    |||

QY 383 TACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTG 442
    |||
DB 359 ATCAGAGGAATTCGTTTCTGGAGGAGAGAGTGCTACATTTAAAGCGCAAGTGAAGGCT 418
    |||
QY 443 ATT-----CCTGAATTTTCTGAACAGAGAGGAAATAGATGAGAAATGAAGAA----- 490
    |||
```

```
Db 419 CGTATTCCTGAGGTGGCGCGGTGACCAACACAGAGCATCTCTCCAACTGGAAGGCAAG 478
    |||
QY 491 ---ATTACCACTTTCTTTTGAACAGTCACTGATTTGGGTCCCGACAGAAAAGCCTATT 547
    |||
DB 479 ATCATGCCAGTCAAAATATGAAGAAAATTCCTTATCTGGGTGGCTGTAGATCAGCCTGTG 538
    |||
QY 548 GAAAACCGAGATTTTCTTAAATAATTCAAAATTTCTGAGATTTTGTGATAACGTGACCATG 607
    |||
DB 539 AAGGACAAACAGCTTCTTG---AATTCTAAGGTGTAGAACTCTGCGGTGACCTTCTTATT 595
    |||
QY 608 TATTGGATCAATCCCACTCTTAATATCATAGTTTCTGAGTTTACAAGACTTTTGAGGAGGAGGA 667
    |||
DB 596 TTCTGGCTTAAACCAACCTATCCAAAAGAAATCCAGAGGGAAGAGAGAGTGGTAAGA 655
    |||
QY 668 GAAGATCTTCACATTTCTGCGCAACGAAAAGGAGGATTTGAAACAAAATGAACAGTGGGTG 727
    |||
DB 656 AAAAATGTTTCCAACTACCAAAAAGACCACACAGTGGACACAGGAGCAACCCAGGGGCT 715
    |||
QY 728 GTCCTCTCAAGTGAAGTAGAGAGACCCGT-----CAGCCAGACAAGCAAGTGAAGAA 781
    |||
DB 716 GGAAGACTGAATTAATGAACACGACCAAGTGTTTCAAGAGGACTCAAGCCTTCAATCCT 775
    |||
QY 782 GAATTTCCAATAATGACTATATCTGAAAATGGAATAGAAATTTGATCCCATGCTGGATGAG 841
    |||
DB 776 GATAATCCTTATCATCAGCAGGAAGGGAAGCATGACATTCGACCTAGACTGGATCAC 835
    |||
QY 842 AGAGGTATTTGTTGTTTACTCTCCGTCGAGGCAACCGCTATTGGCCGCCGCTCTGTGAA 901
    |||
DB 836 GAAGGAATCTGTTGTATAGAAATGTAGCGGAGCTACACCCACTGCCAAGAGATCTGTGAA 895
    |||
QY 902 CCCTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGGAGGAGTCACTCTGTCGT 961
    |||
DB 896 CCCTGGGGGCTATTACCCATGCGCTTATTAATTAATCAAGCTGCCCTTCGGCTTGCGA 955
    |||
QY 962 GTCATCATGCTTTGTAACCTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAA 1012
    |||
DB 956 GTCATCATGCCATGTAGTGGTGGTGGCCCGCTATCTTGGGCATGGTGGA 1006
    |||

RESULT 6
US-08-241-465B-5
; Sequence 5, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI
; APPLICANT: YUJI HIRAKI
; APPLICANT: KAZUHIRO TAKAHASHI
; APPLICANT: JUNKO SUZUKI
; APPLICANT: JUN KONDO
; APPLICANT: ATSUKO KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,465B
; FILING DATE: May 11, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
US-08-241-465B-6
```

Db	896	CCCTGGGGGGCTATTACCATGGCTTTATAAATTATCAAGGCTGCCGCTTCCGCTGCAGA	955
QY	962	GTCAATCATGCTTGTACTGTGGTGGCCCGCATGCTGGGAGGTTCTAA	1012
Db	956	GTCAATCATGCTTGTACTGTGGTGGCCCGCATGCTTGGGCGCATGGTGTGA	1006
<p>RESULT 7</p> <p>US-08-241-465B-4</p> <p>; Sequence 4, Application US/08241465B</p> <p>; Patent No. 5719125</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: FUJIO SUZUKI</p> <p>; APPLICANT: YUJI HIRAKI</p> <p>; APPLICANT: KAZUHIRO TAKAHASHI</p> <p>; APPLICANT: JUNKO SUZUKI</p> <p>; APPLICANT: JUN KONDO</p> <p>; APPLICANT: ATSUKO KOHARA</p> <p>; APPLICANT: AKIKO MORI</p> <p>; APPLICANT: EI YAMADA</p> <p>; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN</p> <p>; NUMBER OF SEQUENCES: 21</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Wenderoth, Lind &amp; Ponack</p> <p>; STREET: 805 Fifteenth Street, N.W., #700</p> <p>; CITY: Washington</p> <p>; COUNTRY: D.C.</p> <p>; ZIP: 20005</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: Patentin Release #1.0, Version #1.25</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/241,465B</p> <p>; FILING DATE: May 11, 1994</p> <p>; CLASSIFICATION: 435</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: Warren M. Cheek, Jr.</p> <p>; REGISTRATION NUMBER: 33,367</p> <p>; REFERENCE/DOCKET NUMBER:</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (202) 371-8850</p> <p>; TELEFAX:</p> <p>; TELEFAX:</p> <p>; INFORMATION FOR SEQ ID NO: 4:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 892 base pairs</p> <p>; TYPE: nucleic acid</p> <p>; STRANDEDNESS: double</p> <p>; TOPOLOGY: linear</p> <p>; ORIGINAL SOURCE:</p> <p>; ORGANISM: Human being (Homo Sapiens)</p> <p>; IMMEDIATE SOURCE:</p> <p>; CLONE: pHCN-16-5</p> <p>; FEATURE:</p> <p>; NAME/KEY: P CDS</p> <p>; LOCATION: 2 .. 889</p> <p>; IDENTIFICATION METHOD: E</p> <p>; US-08-241-465B-4</p>			
<p>Query Match 7.6%; Score 90.8; DB 1; Length 892;</p> <p>Best Local Similarity 59.8%; Pred. No. 2.3e-16;</p> <p>Matches 152; Conservative 0; Mismatches 102; Indels 0; Gaps 0;</p>			
QY	759	ACGCCAGACAAGCAAGTGGAGGAACCTTCCAAATAATGATATCTATCATCAGCAGGAAGGGAAGCATGA	818
Db	639	AGGACTCAACAGCCTTCAATCTCTGATAATCTTATCATCAGCAGGAAGGGAAGCATGA	698
QY	819	AATTTGATCCCATGCTGGATGAGAGGTTATTTGTTATTTACTCCGCTGAGGCAACC	878
Db	699	CATTGACCTAGACTGGATCAGGAAGGAATCTGTTGTATGATATGTTAGCGGAGCTACA	758





```
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 250
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "N is T, G, A or C"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-047-033-5
Query Match 7.5%; Score 90; DB 1; Length 364;
Best Local Similarity 63.8%; Pred. No. 2.5e-16;
Matches 132; Conservative 2; Mismatches 73; Indels 0; Gaps 0;
QY 806 GAAATGGAATGATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTTACTGC 865
Db 158 GGAGAAAGCATGACATTCGACCCGACATGGAAGGATCTGCTGATAGATGC 217
QY 866 CGTCGAGGCAACCGCTATTGCGCGCGCTGTGAACTTTTACTAGGCTACTACCCATAT 925
Db 218 AGGAGGAGCTACACCCACTGCCAGAGATCKNGAGCCTCTGGGGGCTACCCCATCG 277
QY 926 CCATCTGCTACCAAGGAGGAGGTCATCTGCTGTCATCATGCTGCTGTAAGTGG 985
Db 278 CCTATAAATACCAAGGCTGCCGTTCCGCTGCAGAGTCATCATGCTGCTGAGTGG 337
QY 986 GTGCCCCGATGCTGGGGAGGGTCTAA 1012
Db 338 GTGCCCCGATCTCTGGGCATGGTGTA 364
```

RESULT 10  
US-08-047-033-4  
; Sequence 4, Application US/08047033  
; Patent No. 544157  
; GENERAL INFORMATION:  
; APPLICANT: FUJIO SUZUKI et al.  
; TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/047,033  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/821,859  
; FILING DATE:  
; APPLICATION NUMBER: US/07/745,497  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM: bovine  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE: fetal cartilage  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 251  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "N is T, G, A or C"  
; PUBLICATION INFORMATION:  
; AUTHORS:



APPLICANT: Jun KONDO  
APPLICANT: Atsuko KOHARA  
APPLICANT: Akiko MORI  
APPLICANT: Ei YAMADA  
TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
COUNTRY: D.C.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/241,465B  
FILING DATE: May 11, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 892 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Human being (Homo Sapiens)  
FEATURE:  
NAME/KEY: P CDS  
LOCATION: 2 .. 889  
IDENTIFICATION METHOD: E  
US-08-241-465B-7

Query Match 7.2%; Score 86; DB 1; Length 892;  
Best Local Similarity 58.7%; Pred. No. 5.6e-15;  
Matches 149; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 759 AGCCAGACAGCAAGTGGAGGAGAACTTCCAAATAATGACTATCTACTGAAATGGAATAG 818  
Db 639 AGGACTCAAGCCCTTCAATCCTGATAATCTTATCATCAGAGGAGGGGAAAGCATGA 698  
QY 819 AATTTGATCCCATGCTGGATGAGAGAGGTATTTGTTATTTACTGCCCTCGAGGCAACC 878  
Db 699 CATTGACCCCTAGATGGATCAGGAAGAACTGTTGTATAGAAATGAGCGGAGCTACA 758  
QY 879 GCTATTGGCCGCGTCTGTGAACTTTACTAGGCTACTACCATATCCATCTGCTACC 938  
Db 759 CCACCTGCCAGAAGATCTGTGAACCCCTGGGGGGCTATTACCATGGCCTTATAATTATC 818  
QY 939 AAGGAGGAGCAGTCTCTCTGTCATCATCCCTGTAACCTGGTGGTGGCCCGCATGC 998  
Db 819 AAGGCTGGCTGCGCTGCAGAGTCATCATCCATGAGTGGTGGTGGCCCGCTATCT 878  
QY 999 TGGGAGGAGGTCTAA 1012  
Db 879 TGGGATGGTGTGA 892

RESULT 13

US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
QY 795 ATGACTACTGAAATGGAATAGAA 820

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZ9pt-Fls  
US-08-232-463-14

Query Match 4.4%; Score 53.2; DB 1; Length 7218;  
Best Local Similarity 4.1%; Pred. No. 4.5e-05;  
Matches 16; Conservative 216; Mismatches 154; Indels 0; Gaps 0;  
QY 435 TTAAGTGTCTCTGAACAGTCAGTGATTGGGTCCAGCAAAAGCCTATTGAAACC 554  
Db 1459 TTAAGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400  
QY 495 CCACAACCTTCTTGAACAGTCAGTGATTGGGTCCAGCAAAAGCCTATTGAAACC 554  
Db 1399 RRR 1340  
QY 555 GAGATTTCTTTAAATAATCCAAAATCTGGAGATTGTGATAACCTGACCATGTATTGA 614  
Db 1339 RRR 1280  
QY 615 TCAATCCACTCTAATATCAGTTTCTGAGTACAGACTTTGAGGAGGAGGAGAGATC 674  
Db 1279 RRR 1220  
QY 675 TTCCTTTCTGCCAACGAAAAAGGATTGAACAAAATCAACAGTGGTGGTCCCTC 734  
Db 1219 RRR 1160  
QY 735 AAGTGAAGTAGAGAAGACCCGTCACGCCAGACAGCAAGTAGGAGAACTTCCAATAA 794  
Db 1159 RRR 1100  
QY 795 ATGACTACTGAAATGGAATAGAA 820

Db 1099 RRRRRRRRRRRRRRRRRRRRRRRRR 1074

RESULT 14

US-08-047-033-7

Sequence 7, Application US/08047033

Patent No. 5441157

GENERAL INFORMATION:

APPLICANT: FUJIO SUZUKI et al.

TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,033

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/821,859

FILING DATE:

APPLICATION NUMBER: US/07/745,497

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856

TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORGANISM: bovine

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE: fetal cartilage

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-047-033-7

Query Match 4.1%; Score 49.2; DB 1; Length 290;

Best Local Similarity 60.4%; Pred. No. 0.00013;

Matches 81; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 806 GAAATGGAATAGAAATTGATCCCATGCTGGATGAGAGAGGTTATTGTGTATTACTGC 865

Db 157 GGAGAAAGCATGACATTCGACCCCGAGCTGGATCATGAAGGCATCTGCTGTATAGAATGC 216

QY 866 CGTCAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATAT 925

Db 217 AGGAGGAGCTACACCCACTGCCAGAGATCTGTGAGCCTCTGGGGGGCTACCCCATGG 276

QY 926 CCATACCTGCTACCA 939

Db 277 CCCTATACTACCA 290

RESULT 15

US-09-453-702B-14

Sequence 14, Application US/09453702B

Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.

Burland, Valerie

Perna, Nicole T.

Plunkett, Guy

Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady

STREET: 1 South Finckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 12886

TYPE: nucleic acid

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-453-702B-14

Query Match      3.5%; Score 41.8; DB 4; Length 12886;
Best Local Similarity 45.5%; Pred. No. 0.12;
Matches 135; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

QY 415 ATGTTTATCAAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACCGAGAGGAAAT 474
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3283 ATGAATTAAAGAGTGTAACTGAGAGTCGTTAATTTCTTGATATTCCTAGCGAGGAGA 3342
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 475 AGATGAGATGAAGAAATACCAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGC 534
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3343 CAAAAGAAAGTCTGAATTTAAATTCCTTTTAACTCAGAGAGATGATCCTCTTAGC 3402
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 535 AGAAAGCCTATTGAAACCGAGATTTCTTAAAAAATTCAAAAATTCGGAGATTTGTGA 594
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3403 AGAAACCCCAATAAAAAAGCCAGATATTTTAAAAAAATC-CAATTGCCAATATTTATGA 3461
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 595 TAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTT 654
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3462 GGAAATTGAAAGNGCAATTTATTATCTGAAATANGANCNCNNTGNTTCCANAGAACN 3521
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 655 TGAGGAGGAGGGAGAGATCTTCACCTTCTGCGAACGAAAAAAGGGATTGAACA 711
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3522 TNANAAGNANAGCCANATNTTNAATCNCGAANANANATTNAAAGNATTGNTNA 3578
Db ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: September 3, 2004, 06:18:17  
Job time : 85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 07:44:08 ; Search time 2370 Seconds  
(without alignments)  
15082.291 Million cell updates/sec

Title: US-10-063-730-115

Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthu.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_estl.\*

9: gb\_estl.\*

10: gb\_est2.\*

11: gb\_estc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mus.\*

23: em\_gss\_mam.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_pbq.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	965.8	80.7	1204	11 AK003748	AK003748 Mus muscu
2	954	79.7	954	29 AY408390	AY408390 Homo sapi
3	951.4	79.5	954	29 AY408391	AY408391 Pan trogl
4	949	79.3	1159	11 AK014761	AK014761 Mus muscu

5	828.6	69.2	926	13	BQ934226	BQ934226 AGENCOURT
6	826	69.0	954	29	AY408392	AY408392 Mus muscu
7	708.8	59.2	1000	13	BY704009	BY704009 BY704009
8	692.4	57.8	1011	12	BI456138	BI456138 603172901
9	674.8	56.4	800	14	CF410248	CF410248 CH3065 G
10	674	56.3	919	14	CB195636	CB195636 AGENCOURT
11	668.4	55.4	961	12	BI694284	BI694284 603347729
12	660.6	54.8	845	14	CB588736	CB588736 AGENCOURT
13	623.8	52.1	803	14	CF410247	CF410247 CH3065 G
14	623	52.0	932	13	BY714257	BY714257 BY714257
15	620.6	51.8	876	14	CB587595	CB587595 AGENCOURT
16	612.6	51.2	751	14	CD512100	CD512100 AGENCOURT
17	608.2	50.8	867	14	CB587917	CB587917 AGENCOURT
18	579.6	48.4	744	12	BI454468	BI454468 603171428
19	579	48.4	846	10	BF182340	BF182340 601804094
20	577	48.2	958	12	BGI74435	BGI74435 602334255
21	574	48.0	712	14	CD103859	CD103859 AGENCOURT
22	553.2	46.2	679	12	BI248748	BI248748 602992715
23	531.8	44.4	600	9	AV593197	AV593197 AV593197
24	530.4	44.3	639	14	CD773806	CD773806 AGENCOURT
25	500.6	41.8	735	12	BI694698	BI694698 603347638
26	498.6	41.7	557	12	BI535437	BI535437 398969 MA
27	494	41.3	588	10	AW743952	AW743952 ur24h02.Y
28	493.6	41.2	798	12	BI249999	BI249999 602995241
29	485.8	40.6	610	14	CB440637	CB440637 6090932 MA
30	485.4	40.6	636	12	BI304069	BI304069 UI-R-DR0-
31	474.8	39.7	479	9	AI123839	AI123839 qa74e02.x
32	473	39.5	595	13	BI440284	BI440284 690548 MA
33	468	39.1	468	13	BI104255	BI104255 BX104255
34	463.6	38.7	888	10	BF121576	BF121576 601756977
35	460	38.4	462	10	BF439260	BF439260 nab61h07.
36	435.8	36.4	465	14	T12179	T12179 A533F Heart
37	424.6	35.5	856	10	BF123957	BF123957 601759641
38	419	35.0	741	14	CB317710	CB317710 AGENCOURT
39	409.4	34.2	521	13	BX638501	BX638501 BX638501
40	385	32.2	399	9	AI039039	AI039039 ok30b10.8
41	379.8	31.7	1224	14	CB202990	CB202990 AGENCOURT
42	378.6	31.6	733	10	BE375361	BE375361 601230345
43	376.4	31.4	460	9	AI146280	AI146280 qb87b10.x
44	371.6	31.0	468	9	AV593196	AV593196 AV593196
45	369	30.8	433	14	CF112025	CF112025 Shultzomi
46	366.8	30.6	487	9	AI112003	AI112003 UI-R-YO-m
47	366.2	30.6	504	10	BF523627	BF523627 UI-R-Cl-1
48	360.8	30.1	404	12	BI535434	BI535434 398963 MA
49	358.8	30.0	452	9	AI502787	AI502787 UI-R-Cl-1
50	358.4	29.9	882	10	BF679243	BF679243 602153429
51	352.2	29.4	509	14	CD572281	CD572281 PBL 18 EO
52	350.4	29.3	407	10	BF150118	BF150118 uy82a03.Y
53	336.8	28.1	359	9	AA236166	AA236166 z842f09.8
54	332.6	27.8	391	10	AW962409	AW962409 EST374482
55	329.6	27.5	419	14	W97621	W97621 mf97e12.r1
56	327.6	27.4	416	13	BY377460	BY377460 BY377460
57	319.2	26.7	434	10	AW144807	AW144807 EST291897
58	312	26.1	452	10	BH831558	BH831558 BH831558
59	309.6	25.9	661	10	BE288641	BE288641 601094042
60	308.4	25.8	337	9	AA297231	AA297231 EST112774
61	307	25.6	424	9	AI600132	AI600132 EST251835
62	306.8	25.6	387	13	BY772480	BY772480 BY772480
63	298.8	25.0	392	13	BY378738	BY378738 BY378738
64	298	24.9	299	9	AI147044	AI147044 gb82d04.x
65	297.4	24.8	487	13	BX517933	BX517933 BX517933
66	293.8	24.5	380	10	BF148505	BF148505 uy82a09.x
67	285	23.8	448	13	BY452021	BY452021 BY452021
68	284.8	23.8	860	13	BUI27029	BUI27029 603113730
69	274	22.9	369	10	AW240909	AW240909 uq38a03.x
70	269.8	22.5	542	14	CB114632	CB114632 K-EST0158
71	269.6	22.5	345	13	BY112564	BY112564 BY112564
72	268.8	22.5	349	13	BY778741	BY778741 BY778741
73	268.6	22.4	352	13	BY111994	BY111994 BY111994
74	266.6	22.3	342	13	BY115292	BY115292 BY115292
75	266	22.2	1109	13	BUI131955	BUI131955 603119028
76	263.4	22.0	338	13	BY789211	BY789211 BY789211
77	261.6	21.9	341	13	BY086042	BY086042 BY086042





```
Db      98  CCTCAGAGAACTGTGAGGCTGTACATCTTAAATGAGAAAGCTCTGAAATCTAAGAAG 157
Qy      131 ATATGTAATCACTAAGATTGTGGAATGTGTGTTTGGTATCTCTGGCCCTAACTCTAATT 190
Db      158 ATATGTAATCACTGAAGATTGTGGAATGTGTGTTTGGTATCTCTGGCCCTTAACITCTAATT 217
Qy      191 GTCCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAAGCCCTATGACATG 250
Db      218 GTCCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTATCCAAGAAAAACCTATGACATG 277
Qy      251 GAGCACACTTCTACAGCAATGAGAGAGAGAGATTATCATGGAAATGATCCTGTG 310
Db      278 GAGCACACTTCTACAGCAATGAGAGAGAGAGATTATCATGGAAATGATCCTGTG 337
Qy      311 ACCAGAACTGAAATATTTCAGAAAGCGGAAATGGCACTGATGAAACATTTGGAAGTGACAC 370
Db      338 ACCAGAACTGAAATATTTCAGAAAGTGGAATGGCACTGATGAAACATTTGGAAGTGACAC 397
Qy      371 TTTAAAAACGATACACTGGGATCTACTTCTGGGTCTTCAAAAATGTTTATCAAAACT 430
Db      398 TTTAAAAATGGATACACTGGGATCTACTTCTGGGTCTTCAAAAATGTTTATCAAAACT 457
Qy      431 CAGATTAAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGATGAGAAATGAAGAA 490
Db      458 CAAATCAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGATGAGAAATGAAGAA 515
Qy      491 ATTACCAAACTTTCTTTTGAACAGTCACTGATTTGGGTCTCCAGCAGAAAAAGCCCTATTGAA 550
Db      516 ATTACTACAACTTTCTTTTGAACAGTCACTGATTTGGGTCTCCAGCAGAAAAAGCCCTATTGAA 575
Qy      551 AACCGAGATTTCTTTAAATTCGAAATTTCTGGAGATTTGTGATAACGTGACCATGTAT 610
Db      576 AACAGAGACTTCTGAAATTTCTGAAATTTCTGGAGATTTGTGATAACGTGACCATGTAT 635
Qy      611 TGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCTGAGATTTGTGATAACGTGACCATGTAT 670
Db      636 TGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCTGAGATTTGTGATAACGTGACCATGTAT 695
Qy      671 GATCTTCACTTTCTCTGCAACGAAAAAGGGATTGAAACAAATGAACAGTGGGTGTC 730
Db      696 GATCTTCACTTTCTCTGCAACGAAAAAGGGATTGAAACAAATGAACAGTGGGTGTC 755
Qy      731 CCTCAAGTGAAGTAGAGAGACCGGTACCGCAGACAAGCAAGTAGAGAGAACTTTCCA 790
Db      756 CCGCAAGTAGAGTAGAGAGACCGGTACCGCAGACAAGCAAGTAGAGAGAACTTTCCA 815
Qy      791 ATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTAT 850
Db      816 ATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTAT 875
Qy      851 TGTGTATTACTGCGTGAGGCAACCGCTATTGCGCGCGCTCTGTGAACCTTTACTA 910
Db      876 TGTGTATTACTGCGTGAGGCAACCGCTATTGCGCGCGCTCTGTGAACCTTTACTA 935
Qy      911 GGTACTACCCATATCCATCTCTACAAAGAGAGCGAGTCACTGTGTCGTGTCATCATG 970
Db      936 GGTACTACCCATATCCATCTCTACAAAGAGAGCGAGTCACTGTGTCGTGTCATCATG 995
Qy      971 CTTTGTAATCTGGTGGGTGGCCGCGATGCTGGGAGAGGTCTTAATAGAGGTTTGAGCTCAA 1030
Db      996 CTTTGTAATCTGGTGGGTGGCCGCGATGCTGGGAGAGGTCTTAATAGAGGTTTGAGCTCAA 1055
Qy      1031 ATGCTTAAATCTGCT---GGCAACATATAATATGATGCTATTCAATCAATTTCTGCG 1086
Db      1056 ACCTTTAACTTCTGTTAGCAATATATTAATTAATGATGCTACTCTCAATTAATTTCTGCC 1115
Qy      1087 TATGAGGATCTGGCCCTCTGGTAGCCAGCTCTCCAGAAATTTACTTGTAGTAAATTCCTCTC 1146
Db      1116 TATGAGGATCTGGCCCTCTGGTAGCCAGCTCTCCAGAAATTTACTTGTAGGATATTCCTCTC 1175
Qy      1147 TTCATGTTCTTAATAACTTCTACATTTATC 1175
Db      1176 TTCATGTTCTTAATAACTTCTACATTTATC 1204
```

RESULT 2  
AY408390  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

TITLE  
JOURNAL

COMMENT

FEATURES  
source

gene

ORIGIN

Query Match 79.7%; Score 954; DB 29; Length 954;  
Best Local Similarity 100.0%; Pred. No. 7.9e-173;  
Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 ATGGCAAGAAATCTCTCCAGAGAAATGTGAAGACTGCACATTTCTAAATGCAGAGCTTTT 118  
Db 1 ATGGCAAGAAATCTCTCCAGAGAAATGTGAAGACTGCACATTTCTAAATGCAGAGCTTTT 60  
Qy 119 AAATCCAAGAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGTTATCCTGGCC 178  
Db 61 AAATCCAAGAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGTTATCCTGGCC 120  
Qy 179 CTAACCTCTAATTGTCTCTGTTTGGGGGAGCAAGCTTCTGGCGGAGGTACCCAAAAA 238  
Db 121 CTAACCTCTAATTGTCTCTGTTTGGGGGAGCAAGCTTCTGGCGGAGGTACCCAAAAA 180  
Qy 239 GCCTATGACATGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAA 298  
Db 181 GCCTATGACATGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAA 240  
Qy 299 ATGTATCTGTGACAGAACTGAAATATTTCAGAGCGGAAATGGCACTGATGAACATTTG 358  
Db 241 ATGTATCTGTGACAGAACTGAAATATTTCAGAGCGGAAATGGCACTGATGAACATTTG 300  
Qy 359 GAAGTGCAGCTTTTAAACCGATACACTGCGCATCTACTCTGGGTCTTCAAAATGT 418  
Db 301 GAAGTGCAGCTTTTAAACCGATACACTGCGCATCTACTCTGGGTCTTCAAAATGT 360  
Qy 419 TTTATCAAACTCAGATTAAAGTGATTCTCTGAATTTCTGAACCCAGAGAGAAATAGAT 478  
Db 361 TTTATCAAACTCAGATTAAAGTGATTCTCTGAATTTCTGAACCCAGAGAGAAATAGAT 420

AY408390 954 bp DNA linear GSS 12-DEC-2003  
Homo sapiens HCM3208 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

AY408390  
AY408390.1 GI:39764361  
GSS.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 954)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 954)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Adams,M.D. and Cargill,M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering  
them based on alignment.

Location/Qualifiers

1..954

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

<1..>954

/locus\_tag="HCM3208"

```
QY 479 GAGATGAGAAATTACACAACTTCTTTGAAAGTCAGTGATTTGGTCCCGACAGAA 538
DB 421 GAGAAATGAAGAAATTTACACAACTTCTTTGAAAGTCAGTGATTTGGTCCCGACAGAA 480
QY 539 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTTGTGATAAC 598
DB 481 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTTGTGATAAC 540
QY 599 GTGACCATGTATTTGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 658
DB 541 GTGACCATGTATTTGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 600
QY 659 GAGAGGGAGAGATCTTCACTTTCTCCAGCAAGAAAAAGGATTTGAACAAATGAA 718
DB 601 GAGAGGGAGAGATCTTCACTTTCTCCAGCAAGAAAAAGGATTTGAACAAATGAA 660
QY 719 CAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGCTCAGCCAGCAAGCAAGTGAAG 778
DB 661 CAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGCTCAGCCAGCAAGCAAGTGAAG 720
QY 779 GAAGAACTTCCAAATAATGACTATCTGAAATGGAATGGAATTTGATCCCACTGCGAT 838
DB 721 GAAGAACTTCCAAATAATGACTATCTGAAATGGAATGGAATTTGATCCCACTGCGAT 780
QY 839 GAGAGAGGTATTGTGTATTTACTGCGTGCAGGCAACCGCTATTGCGCGCGCTCTGT 898
DB 781 GAGAGAGGTATTGTGTATTTACTGCGTGCAGGCAACCGCTATTGCGCGCGCTCTGT 840
QY 899 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 958
DB 841 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 900
QY 959 CCGTGCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012
DB 901 CCGTGCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 954

RESULT 3
AY408391
LOCUS
DEFINITION
Pan troglodytes HCM3208 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY408391
VERSION
AY408391.1 GI:39764362
KEYWORDS
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
REFERENCE
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 954)
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..954
/organism="Pan troglodytes"
/mol_type="genomic DNA"
```

```
/db_xref="taxon:9598"
<1..>954
/locus_tag="HCM3208"

gene
ORIGIN

Query Match 79.5%; Score 951.4; DB 29; Length 954;
Best Local Similarity 99.8%; Pred. No. 2.5e-172; Indels 0; Gaps 0;
Matches 952; Conservative 0; Mismatches 2;

QY 59 ATGCCAAAGAAATCTCCAGAGAAATGTGAAGACTGTACATTTCTAAATGCAAGACTTTT 118
DB 1 ATGCCAAAGAAATCTCCAGAGAAATGTGAAGACTGTACATTTCTAAATGCAAGACTTTT 60
QY 119 AAATCCAAAGAAATATATGTAATATCACTTAAATTTGTGACCTGGTGTGGTATCTGCCC 178
DB 61 AAATCCAAAGAAATATATGTAATATCACTTAAATTTGTGACCTGGTGTGGTATCTGCCC 120
QY 179 CTAATCTAAATGTCCTGTTTGGGGAGAGCACTTCTGCGCGAGTACCCAAAAA 238
DB 121 CTAATCTAAATGTCCTGTTTGGGGAGAGCACTTCTGCGCGAGTACCCAAAAA 180
QY 239 GCCTATGACATGGAGCACACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGAA 298
DB 181 NCCTATGACATGGAGCACACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGAA 240
QY 299 ATTGATCCTGTGACAGAACTGAAATATTCAGAACCGGAATGSCACTGATGAAACATG 358
DB 241 ATTGATCCTGTGACAGAACTGAAATATTCAGAACCGGAATGSCACTGATGAAACATG 300
QY 359 GAGTGCAGCACTTTAAATAACGGATACACTGGCATCTTCTGCTGGTCTTCAAAAATGT 418
DB 301 GAGTGCAGCACTTTAAATAACGGATACACTGGCATCTTCTGCTGGTCTTCAAAAATGT 360
QY 419 TTTATCAAAACTCAGATTTAAAGTGAATTTCTGAAATTTCTGAAACCAAGAGAGAAATGAT 478
DB 361 TTTATCAAAACTCAGATTTAAAGTGAATTTCTGAAATTTCTGAAACCAAGAGAGAAATGAT 420
QY 479 GAGAAATGAAGAAATTAACCAAACTTTCTTTGAACAGTCACTGATTTGGTCCCGACAGAA 538
DB 421 GAGAAATGAAGAAATTAACCAAACTTTCTTTGAACAGTCACTGATTTGGTCCCGACAGAA 480
QY 539 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTTGTGATAAC 598
DB 481 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTTGTGATAAC 540
QY 599 GTGACCATGTATTTGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 658
DB 541 GTGACCATGTATTTGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 600
QY 659 GAGAGGGAGAGATCTTCACTTTCTCCAGCAAGAAAAAGGATTTGAACAAATGAA 718
DB 601 GAGAGGGAGAGATCTTCACTTTCTCCAGCAAGAAAAAGGATTTGAACAAATGAA 660
QY 719 CAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGTCAGCCAGCAAGCAAGTGAAG 778
DB 661 CAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGTCAGCCAGCAAGCAAGTGAAG 720
QY 779 GAAGAACTTCCAAATAATGACTATCTGAAATGGAATGGAATTTGATCCCACTGCGAT 838
DB 721 GAAGAACTTCCAAATAATGACTATCTGAAATGGAATGGAATTTGATCCCACTGCGAT 780
QY 839 GAGAGAGGTATTGTGTATTTACTGCGTGCAGGCAACCGCTATTGCGCGCGCTCTGT 898
DB 781 GAGAGAGGTATTGTGTATTTACTGCGTGCAGGCAACCGCTATTGCGCGCGCTCTGT 840
QY 899 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 958
DB 841 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 900
QY 959 CCGTGCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012
DB 901 CCGTGCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 954
```

RESULT 4  
 AK014761  
 LOCUS  
 DEFINITION  
 Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4833424007 product:tenomodulin, full insert sequence.  
 AK014761  
 ACCESSION  
 AK014761.1 GI:12852797  
 VERSION  
 HTG; CAP trapper.  
 KEYWORDS  
 Mus musculus (house mouse)  
 SOURCE  
 Mus musculus  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE  
 High-efficiency full-length cDNA cloning  
 JOURNAL  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636  
 REFERENCE  
 2  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 20499374  
 PUBMED  
 11042159  
 REFERENCE  
 3  
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE  
 20530913  
 PUBMED  
 11076861  
 REFERENCE  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE  
 Functional annotation of a full-length mouse cDNA collection  
 JOURNAL  
 Nature 409, 685-690 (2001)  
 MEDLINE  
 11706861  
 REFERENCE  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL  
 Nature 420, 563-573 (2002)  
 MEDLINE  
 12042159  
 PUBMED  
 12042159  
 REFERENCE  
 6  
 Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirakawa,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,K., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,K., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

## COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT-3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer/adaptor of sequence [5'-GAGAGAGATCTCCAGTAAATAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

## FEATURES

Location/Qualifiers  
 source  
 1..1159  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:4833424007"  
 /db\_xref="MGI:1896631"  
 /db\_xref="taxon:10090"  
 /clone="4833424007"  
 /tissue\_type="head"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="0 day neonate"  
 585..992  
 /note="unnamed protein product; putative tenomodulin (MGI:MGI:1929885, GB|NM\_022322, evidence: BLASTN, 99%, match=1151)"  
 /codon\_start=1  
 /protein\_id="BAC25447.1"  
 /db\_xref="GI:26380843"  
 /translation="MTWNPITLAVSELQDFEEDGDLHPTSEKKGIDNEQWVVPQGVYKPYCYQGGRRVICRVIMPNCNMMVARMGLRV"  
 1141..1146  
 /note="putative"  
 1159  
 /note="putative"

polyA\_signal

polyA\_site

## ORIGIN

Query Match 79.3%; Score 949; DB 11; Length 1159;  
 Best Local Similarity 89.3%; Pred. No. 6.7e-172;  
 Matches 1035; Conservative 0; Mismatches 120; Indels 4; Gaps 1;  
 QY 21 TCTCAAAGCAAGCAAGAGAGTACTGTGCTGAGAGACCATGGCAAGATCTCTCCAGAGA 80  
 Db 1 TCTCAAAGCTGGCAAGAGACCGTGTCTGGGAGACCATGGCAAGATCTCTCCAGAGA 60  
 QY 81 ATGTGCAAGACTGTCACTTAAATGCAAGAGCTTTTAAATCCAGAAATATGTAAT 140  
 Db 61 ACTGTGAGGGCTGTCACTTAAATGCAAGAGCTCTGATATCTAAGATGATATTAAT 120  
 QY 141 CACTTAAGATTTGTGGAGCTGGTGTGGTATCTGGCCCTAACTCTAATTTGCTCTTTT 200  
 Db 121 TACTGAAGATTTGTGGAGCTAGTGTGGTATCTGGCCCTTAATCTAATTTGCTCTTTT 180  
 QY 201 GGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAGCCCTATGACATGGAGCACATT 260  
 Db 181 GGGGAGCAAGCACTTCTGGCCCGATGTATCCAAGATACCTATGATATGGAGCACATT 240  
 QY 261 TCTCAGCAATCGAGAGAGAGAGATTTATCGAATTCATCTGTGACAGACTG 320  
 Db 241 TCTCAGCAATCGAGAGAGAGAGATTTATCGAATTCATCTGTGACAGACTG 300  
 QY 321 AAATATTGAGAGCGGAATGGCACTGATGAACATTTGGAAGTCACGACTTTAAACAG 380  
 Db 301 AAATATTGAGAGCGGAATGGCACTGATGAACATTTGGAAGTCACGACTTTAAATG 360

JOURNAL COMMENT  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M13614 row: 1 column: 09  
High quality sequence stop: 640.  
Location/Qualifiers  
1..926  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6200216"  
/sex="male"  
/tissue\_type="sciatic nerve"  
/dev\_host="adult, 70 Yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupski sciatic nerve"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: Not1; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

FEATURES source  
Query Match 69.2%; Score 828.6; DB 13; Length 926;  
Best Local Similarity 95.7%; Pred. No. 8.2e-149;  
Matches 864; Conservative 0; Mismatches 34; Indels 5; Gaps 1;

ORIGIN  
97 CATTCTAAATGCAGAGCTTTTAAATCCAAGAAAATATCTAAATCACTTAAGATTGTGG 156  
1 CATTCTAAATGCAGAGCTTTTAAATCCAAGAAAATATCTAAATCACTTAAGATTGTGG 60  
157 ACTGGTGTGTGTATCTGCGCCCTAACTTAATTCCTCTGTTTGGGGGAGCAAGCATT 216  
61 ACTGGTGTGTGTATCTGCGCCCTAACTTAATTCCTCTGTTTGGGGGAGCAAGCATT 120  
217 CTGGCCGAGGTACCCAAAAGCCCTATGACATGGAGCACACTTTCTACAGCAATGGAGA 276  
121 CTGGCCGAGGTACCCAAAAGCCCTATGACATGGAGCACACTTTCTACAGCAATGGAGA 180  
277 GAAGAAGAGATTTCATCGAAATTCATCTGTGACAGAACTGAAATATTCAGAGCGG 336  
181 GAAGAAGAGATTTCATCGAAATTCATCTGTGACAGAACTGAAATATTCAGAGCGG 240  
337 AATGGCACTGATGAACATTCGAGTGCACGACTTTTAAACCGGATACACTGGCATCTA 396  
241 AATGGCACTGATGAACATTCGAGTGCACGACTTTTAAACCGGATACACTGGCATCTA 300  
397 CTTCTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTATTCTCTGAATTTTC 456  
301 CTTCTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTATTCTCTGAATTTTC 360  
457 TGAACAGAGAGAAATAGATGAGAAATGAGAAATTCACCAACTTTCTTTGAACAGTC 516  
361 TGAACAGAGAGAAATAGATGAGAAATGAGAAATTCACCAACTTTCTTTGAACAGTC 420  
517 AGTGATTGGTCCCGACGAGAAAGCTTATGAAACCGAGATTCTTAAATAATTCCAA 576  
421 AGTGATTGGTCCCGACGAGAAAGCTTATGAAACCGAGATTCTTAAATAATTCCAA 480  
577 AATCTCGAGATTGTGATAACCGTGCATGTATTGGATCAATCCCACTCTAATATCAGT 636

381 GATACACTGGCATCTACTTCTGTTGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAG 440  
361 GATACACTGGCATCTACTTCTGTTGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAG 420  
441 TGAATCTCTGAATTTCTGAACAGAGAGGAAATAGATGAGAAATGAGAAATATTCACAA 500  
421 TGAATCTCTGAATTTCTGAACAGAGAGGAAATAGATGAGAAATGAGAAATATTCACAA 480  
501 CTTTCTTTGAACAGTCACTGATTTGGTCCCGACGAGAAAGCCCTATTCAGAAACCGAGATT 560  
481 CTTTCTTTGAACAGTCACTGATTTGGTCCCGACGAGAAAGCCCTATTCAGAAACCGAGATT 540  
561 TTTCTTAAATAATTCAAAATCTCGAGATTTGFGATAACCTGACCATGATTTGATGATCAATC 620  
541 TCTCTGAAAATTTCAAAATCTCGAGATTTGCGATAATGTGACCATGATTTGATGATCAATC 600  
621 CCACTCTAATATCACTTTCTGAGTTTCAAGACTTTGAGAGAGGAGGAGAGATCTTCACT 680  
601 CCACTCTAATAGCTTTTCAAGATTTACAGGACTTTGAGAGAGGAGGAGAGATCTTCACT 660  
681 TTTCTGCGCAACGAAAAAAGGGATTGACAAAATGAACAGTGGTGGTCCCTCAAGTGA 740  
661 TTTCTACCAAGTAAAAAAGGGATTGACAAAATGAACAGTGGTGGTCCCTCAAGTGA 720  
741 AAGTAGAAGACCCGTCACGCGCAGACAGCAAGTGAAGAACTTCCAAATAAATGACT 800  
721 AGTGTGAGAGACCCGCCACACAGCAAGCAAGCGAGGAGACCTTCTATAAATGACT 780  
801 ATACTGAAAATGGAATAGAAATTTGATCCCATGCTGATGAGAGAGTTATTTGTTGATTT 860  
781 ATACTGAAAATGGAATTTGAAATTTGACCAATGCTGATGAGAGAGGTTACTTGTGATTT 840  
861 ACTGCGCTGAGGCAACGCTATTGCGCGCGTCTGTGAACCTTTACTAGGCTACTACC 920  
841 ACTGCGCTGAGGCAACGCTATTGCGCGCGTCTGTGAACCTTTACTAGGCTACTACC 900  
921 CATATCCATCTACTGCTACCAAGAGGAGACGAGTCACTCTGCTGTCATFCACTGCTTTGTAAT 980  
901 CATACCCCTACTGCTACCAAGAGGTCGAGTCACTCTGCTGTCATFCACTGCTTTGTAAT 960  
981 GGTGGTGGCCGCGATGCTGTTGGGAGGCTTAATAGAGAGTTTGAAGTCAAAATGCTTTAAAC 1040  
961 GGTGGTGGCCGCGATGCTGTTGGGAGGCTTAATAGAGAGTTTGAAGTCAAAATGCTTTAAAC 1020  
1041 TGCT----GGCAACATATATAATGCTATTAATCAAGATTTCTGCTATGAGGCAAT 1096  
1021 TTTCTGTAGCCAAATATAATTAATGCTATTAATCAAGATTTCTGCTATGAGGCAAT 1080  
1097 CTGGCCCTGGTAGCGCTCTCCAGAAATTAATGTTAGGTAATTCCTCTCTTCATGTTCT 1156  
1081 TTGCTTCAAGTAGCTATCTTCAAGATTTACTTTAGGATATTTCTCTCTTCTCATGTTCT 1140  
1157 AATAAATCTTCAATATTC 1175  
1141 AATAAATCTTCAATATTC 1159

RESULT 5  
BQ934226 926 bp mRNA linear EST 21-AUG-2002  
LOCUS AGENCOURT.8883116 Lupski\_sciatic\_nerve Homo sapiens cDNA clone  
DEFINITION IMAGE:6200216 5', mRNA sequence.  
ACCESSION BQ934226  
VERSION BQ934226.1 GI:22349609  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 926)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```
Db 481 AATTCTGGAGATTGTGATAAAGTGAACCAATGATTTGGATCAATCCCACTCTAATACAT 540
QY 637 TTCTGAGTTTACAGACTTTGAGGAGGAGGAGAGAGATCTTCACTTTCTCTGCAACGAAAA 696
Db 541 TTCTGAGTTTACAGACTTTGAGGAGGAGGAGAGAGATCTTCACTTTCTCTGCAACGAAAA 600
QY 697 AAAAGGGATTGAACAAATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCG 756
Db 601 AAAAGGGATTGAACAAATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCG 660
QY 757 TCACGCCAGACAGCAAGTGAAGAGAACTTCCAATAAATGACTATCTAGAAAATGGAAT 816
Db 661 TCACGCCAGACAGCAAGTGAAGAGAACTTCCAATAAATGACTATCTAGAAAATGGAAT 720
QY 817 AGAATTTGATCCATGCTGGATGAGAGAGATTATTTGTTGTTATTTCTGCGTCGAGGCAA 876
Db 721 AGAATTTGATCCATGCTGGATGAGAGAGATTATTTGTTGTTATTTCTGCGTCGAGGCAA 780
QY 877 CGCTATTTCGCCCGCGTCTGTGAACCTTTACTAGGCTACTACCATATCCATATCTG--- 933
Db 781 CGCTATTTCGCCCGCGTCTGTGAACCTTTACTAGGCTACTACCATATCCATATCTGCGCT 840
QY 934 --CTACCAAGGAGGAGAGTCACTGCTGTCATCATGCTTGTAACTGGTGGTGGCC 991
Db 841 ACCAAGGAGGAGGAGTCACTGCTGCGCGGAGCATGCTTGGGAACTTGGGGGGGG 900
QY 992 GCG 994
Db 901 GGC 903

RESULT 6
AY408392
LOCUS
DEFINITION
Mus musculus HCM3208 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY408392
VERSION
AY408392.1
KEYWORDS
GI:39764363
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 954)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 954)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..954
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..954
/locus_tag="HCM3208"

gene
ORIGIN
Query Match 69.0%; Score 826; DB 29; Length 954;
Best Local Similarity 91.8%; Pred. No. 2.6e-148;
```

```
Matches 874; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 59 ATGGCAAGATCTCCAGAGAAATGTGAAGACTGTACATTTCTAATGCAGAAGCTTTT 118
Db 1 ATGGCAAGATCTCCAGAGAACTGTGAGGCTGTACATTTCTAATGCAGAAGCTCTG 60
QY 119 AAATCCAGAAATATGTAAATCACTTAAGATTTCTGCACTGGTGTGGTATCCTGGCC 178
Db 61 AAATCTAAGAGATATGAAATCACTGAAGATTTGTGAAGTGTGGTATCCTGGCC 120
QY 179 CTAACTCTAAATGCTCTGTTTGGGGGAGCAAGCACTTCTGCGCGAGGTACCCAAAAA 238
Db 121 TTAACCTCTAAATGCTCTGTTTGGGGGAGCAAGCACTTCTGCGCGAGGTATCCAGAAA 180
QY 239 GCCTATGACATGGAGCACTTTCTACAGCAATGAGAGAGAAAGATTTACATGAAA 298
Db 181 ACCTATGACATGGAGCACTTTCTACAGCAACGCGGAGAGAAAGATTTACATGAAA 240
QY 299 ATTGATCTGTGACCAAGCACTGAAATATTCAGACGGGAATGGCACTGATGAACATG 358
Db 241 ATTGATCCCAATACAGCAAGCAAGATATTCAGAGTGGAAATGGCACTGATGAACATG 300
QY 359 GAAGTCACGACTTTTAAACCGGATACACTGCGATCTACTTCTGCGGTCTTCAAAAAATG 418
Db 301 GAAGTCATGACTTTTAAACATGGATACACTGCGATCTACTTCTGAGGTCTTCAAAAAATG 360
QY 419 TTTATCAAACTCAGATTAAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGAT 478
Db 361 TTTATTAACACTCAATCAAAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGAT 420
QY 479 GAGATGAAGAAATTAACCAACTTTCTTTGAAACAGTCAGTGTATTTGGTCCAGAGAA 538
Db 421 GAGATGAAGAAATTAACCAACTTTCTTTGAAACAGTCAGTGTATTTGGTCCAGAGAA 480
QY 539 AAGCTATTGAAAAACGAGATTTTCTTAAATAATTCGAAATTTCTGAGATTTGTGATAAC 598
Db 481 AAGCTATTGAAAAACGAGATTTTCTTAAATAATTCGAAATTTCTGAGATTTGTGATAAC 540
QY 599 GTGACCATGATTTGGATCAATCCCACTCTTAATATCAGTTCGTAGTTTACAGACTTTGAG 658
Db 541 GTGACCATGATTTGGATCAATCCCACTCTTAATATCAGTTCGTAGTTTACAGACTTTGAG 600
QY 659 GAGGAGGAGAGAGATCTTCACTTTCTGCAACCAAAAAAGAGATTTGAACAAATGAA 718
Db 601 GAGGAGGAGAGAGATCTTCACTTTCTGCAACCAAAAAAGAGATTTGAACCAATGAA 660
QY 719 CAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCGCTCACGCCAGACAGCAAGTAGAG 778
Db 661 CAATGGGTGGTCCCGCAAGTGAAGTAGAAGACCGCTCACGCCAGACAGCAAGTAGAG 720
QY 779 GAAGAACTTCCAATAATGACTATCTAGAAAATGGAATAGAAATTTGATCCCATGCTGGAT 838
Db 721 GAAGAACTTCCAATAATGACTATCTAGAAAATGGAATTTGATCCCATGCTGGAT 780
QY 839 GAGAGAGTTTATTTGTTGTTTACTCCCGTCGAGGCAACCGCTATTTGCCCGCGGTCTGT 898
Db 781 GAGAGAGTTTATTTGTTGTTTACTCTGTCGAGGCAACCGCTTACTGCGCGGTCTGT 840
QY 899 GAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTTACCAGGAGGAGGATCTGT 958
Db 841 GAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTTACCAGGAGGAGGATCTGT 900
QY 959 CGTGTCAATGCTTGTAACTGTTGGTGGCGCCGATGCTGGGGAGGGTCTAA 1012
Db 901 CGTGTCAATGCTTGTAACTGTTGGTGGCGCCGATGCTGGGGAGGAGGATCTAA 954

RESULT 7
BY704009
LOCUS
DEFINITION
BY704009 RIKEN full-length enriched, 18-day embryo whole body Mus
musculus cDNA clone 1110017101 5', mRNA sequence.
ACCESSION
BY704009
```







Expressed sequence tags from Canine heart  
Unpublished (2003)  
Other ESTs: CH#065 G01T3  
Contact: George AL  
Division of Genetic Medicine  
Vanderbilt University  
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0615  
Tel: 615 936 2660  
Fax: 615 936 2661  
Email: al.george@vanderbilt.edu  
Insert Length: 1647 Std Error: 0.00  
Seq primer: T7: TAATACGACTCTATAGGG  
High quality sequence start: 39  
High quality sequence stop: 797.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

	1: .800	/organism="Canis familiaris"
source		/mol_type="mRNA"
		/db_xref="taxon:9615"
		/clone="CH3#065_G01"
		/tissue_type="heart"
		/cell_type="heart"
		/dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
		/clone_lib="Canine heart normalized cDNA Library in pBluescript"
		/note="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NoR; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"
ORIGIN		
Query Match	56.4%;	Score 674.8; DB 14; Length 800;
Best Local Similarity	93.7%;	Pred. No. 2.4e-119;
Matches 714; Conservative	0; Mismatches 47; Indels 1; Gaps 1;	
QY	42	CTGTGTCGTGAGACACCATGGCAAGAATCCTCCAGAGAATTGTGAAGACTGTGCACATTC 101
DB	39	CTGTGTCGTGAGAGCATGCGAAGAATCCTCCAGAGAACTGTGAGGACTGTGCACATTC 98
QY	102	TAAATGCAGAGCTTTTAAATCCAGAAAATATGTAATCACTTAAGATTGTGCATGG 161
DB	99	TAAATGCAGAGCTTTTAAATCCAGAGAATATGTAATCACTTAAGATTGTGCATGG 158
QY	162	TGTTTGTTGATCTCGGCCCTACTCTAATTTGTCCTGTTTTGGGGGAGCAAGCACITCTGGC 221
DB	159	TGTTTGTTGATCTCGGCTCTAATTTGTCCTGTTTTGGGGGGGCAAGCACITCTGGC 218
QY	222	CGAGGTACCCAAAAAGCCCTATGACATGGAGCACATTTTCTACAGCAATGGAGAGAAGA 281
DB	219	CGAGACACCCAAAAAACATATGACATGGAGCACATTTTCTACAGCAATGGAGAGAAGA 278
QY	282	AGAAGATTACATGGAATAATGATCTCTGTGACCAAGCTGAAATATTCAGAAGCGGAATG 341
DB	279	AGAAGATTACATGGAATAATGATCTCTGTGACCAAGCTGAAATATTCAGAAGTGGAAATG 338
QY	342	GCAC TGATGAAACATTTGGAAGTGACAGCACTTTTAAAAACGGATACACTGGCATCTACTTCG 401
DB	339	GCACTGATGAAACATTTGGNAGTACATGACTTTTAAAAATGGTTTACATGGCATCTACTTCG 398
QY	402	TGGGTCTTCAAANAATGTTTTTATCAAACA CTGAGATTAAAGTGNATTCCTGAATTTCTGAAC 461
DB	399	TAGGCTTTCAAANAATGCTTT CATCAAACCTCAGATTAAAGTGNATTCCTGAATTTCTGAAC 458
QY	462	CAGAGAGCAATAGATGATGAGATGAAGAAATPACCACAACTTTCTTTGAAACAGTCAGTGA 521
DB	459	CAGAGAGCAATAGATGATGAGATGAAGAAATCACCACAACTTTTTTTTGAACAGTCGTGA 518
QY	522	TTTGGTCCCGACAGAAAAAGCCCTATTGAAAA CCGAGATTTTCTTAAAAATTCCAAATTC 581
DB	519	TTTGGTCCCGACAGAAAAAGCCCTATTGAAAAATCGACATTTCTGAAAAATTCCAAATTC 578

QY	582	TGGAGATTGTGATACCGTGACCATCTATTGGATCAATCCCACTCTTAATATCAGTTTCTG	641
Db	579	TGGAAATTTGTGATATGACCATGTATTTGGATCAATCCCACTCTTAATAGCAGTTTCAG	638
QY	642	AGTTACAAGACTTTGAGGAGGAGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAAAG	701
Db	639	AGTTACAAGACTTTGAGGAGGATGGTGAAGATCTTCACTTTCTGCCAATGAAAAAAAAG	698
QY	702	GGATTGAACAAATAAGAACAGT-GGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCAC	760
Db	699	GTAATTGAACAAACACGACGATGGGGTGGTCCCCCAAGTGAAGATGAAAGAATCCGCAC	758
QY	761	GCAGACACGACGAGTGAAGGAAGAACTTCCAATAAATGACTAT	802
Db	759	ACCAGACAAAGCAAGTGAAGGAAGAACTTCCATAAATGACTAT	800
RESULT 10			
CB195636			
LOCUS	CB195636	919 bp	mRNA linear EST 05-FEB-2003
DEFINITION	AGENCOURT_11259532 NIH_MGC_135 Mus musculus cDNA clone		
IMAGE:	30137293 5', mRNA sequence.		
ACCESSION	CB195636		
VERSION	CB195636.1 GI:28222191		
KEYWORDS	EST,		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 919)		
JOURNAL	NTH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: gcapbs-x@mail.nih.gov Tissue Procurement: Dr. David Rowe cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: NDAM0038 row: e column: 14 High quality sequence stop: 662. Location/Qualifiers		
FEATURES			
source	1..919		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:30137293"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_135"		
	/note=Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI; Normalized full-length enriched library from pooled mouse embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5, and 15.5 (size selected for the 0.5-1 kb fragments) Cloned directionally, priming method: Oligo-dr. cDNA enrichment: >1k bp, Average insert size 1.6k bp. Normalization (Cot value): 7.5 kb. Priming sequence: 5'GACTATGTTCTAGTCAGCGCGCGCCCT(T)3' Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."		

## ORIGIN

	Query Match	56.3%	Score 674	DB 14	Length 919
	Best Local Similarity	89.4%	Pred. No. 3.3e-119		
	Matches 748	Conservative 0	Mismatches 86	Indels 3	Gaps 2
Qy	11	CTCTCAGTCCTCTCAAAAGCAAGGAAGAGTACTGTGTGTCTGAGAGACCATGGCCAAAGAAT	70		
Db	13	CTCTCAGTCCTCTCAAAAGCAAGGAAGAGACACCGTGTGTCTGGGAGACCATGGCCAAAGAAT	72		



```
QY 71 CCTCCAGAGAAATTTGTAAGACTGTCTCACTTAAATGCGAAGCTTTTAAATCCAAAGAAA 130
Db 73 CCTCCAGAGAACTGTGAGGCTGTCACTTAAATGCGAAGCTCTGAAATCTAAGAAG 132
QY 131 ATATGTAATACCTTAAGATTTGTGACTGGTGTGTTGGTATCTGCGCCCTAACTCTAAT 190
Db 133 ATATGTAATACCTTAAGATTTGTGACTAGTGTGTTGGTATCTGCGCCCTAACTCTAAT 192
QY 191 GTCTCTTTTGGGGAGCAAGCACTTCTGCGCGGAGGTACCCAAAAGGCCCTATGACATG 250
Db 193 GTCTCTTTTGGGGAGCAAGCACTTCTGCGCGGAGGTATCCAGAAACCTATGACATG 252
QY 251 GAGCACATCTTCTACGCAATCGAGAGAGAGAGATTTAATGAGAAATGATCCTGTG 310
Db 253 GAGCACATCTTCTACGCAATCGCGAGAGAGAGATTTACATGAAATGATCCCAT 312
QY 311 ACCAGAACTGAAATATTACAGACGCGAAATGCGACTGATGAACATGCGAAGTGACGAC 370
Db 313 ACCAGAACGAATATTGAAAGTGGAATGCGACTGATGAACATGCGAAGTGCCATGAC 372
QY 371 TTTAAAAACGGATACACTGGCATCTACTTCTGCGGTCTTCAAAAATGTTTATCAAACT 430
Db 373 TTTAAAAATGGATACACTGGCATCTACTTCTGAGTCTTCAAAATGCTTTTATAAAAT 432
QY 431 CAGATTAAAGTGAATCTCTGAATTTCTGAACAGAGAGAAATAGATGAGAAATGAAGAA 490
Db 433 CAAATCAAAAGTGATTCCTGAATTTCTGAACAGAGAGAAATAGATGAGAAATGAAGAA 492
QY 491 ATTACCAACATCTTCTTGAACAGTCAGTGATTTGGGTCTTCAAAAATGTTTATCAAACT 550
Db 493 ATTACTACAATCTTCTTGAACAGTCAGTGATTTGGGTCTTCAAAAATGCTTTTATAAAAT 552
QY 551 AACCGAGATTTCTTAAAAATTTCAAAAATTTCTGGAGATTTGTGATTAACGTGACCATGTAT 610
Db 553 AACAGAGACTTCTTGAAAAATTTCAAAAATTTCTGGAGATTTGGGATTAATGTGACCATGTAT 612
QY 611 TGGATCAATCCCATCTTAATATAGTTTCTGAGTTAACAAGCTTTGAGGAGGAGGAGAA 670
Db 613 TGGATCAATCCCATCTTAATATAGCTTTTCAAAAATTTACAGGACTTTGAGGAGGAGGAGAA 672
QY 671 GATCTTCACTTCTGCGCAACGAAAAAAGGGATTTGAACTTGAACAAATGAACTGGGTGTC 730
Db 673 GATCTTCACTTCTGCGCAACGAAAAAAGGGATTTGAACTTGAACAAATGAACTGGGTGTC 732
QY 731 CCTCAAGTGAAGTAG-AGAAGACCCGTCAAGCCAGCAAGCAAGTGAAGGAGAACTTCC 789
Db 733 CCGCAAGTGAAGTAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 790 AATTAATGACTTAACTGAAATGG--AATGAAATTTGATCCATGCTGGATGAGAGA 844
Db 793 CTATATGACTTAACTGAAATGGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAAT 849
```

```
RESULT 11
LOCUS BI694284
DEFINITION BI694284 961 bp mRNA linear EST 18-SEP-2001
mRNA sequence.
ACCESSION BI694284
VERSION BI694284.1
KEYWORDS GI:15656913
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 961)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
```

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11955 row: j column: 08  
High quality sequence stop: 892.

FEATURES  
source

1..961  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5375335"  
/issue\_type="tumor, biopsy sample"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam2"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 55.8%; Score 668.4; DB 12; Length 961;  
Best Local Similarity 89.5%; Pred. No. 3.9e-118;  
Matches 764; Conservative 0; Mismatches 86; Indels 4; Gaps 4;

QY 240 CCTATGACATGAGCAGACACTTCTACAGCAATGGAGAGAGAAAGATTTACATGGAAA 299  
Db 41 CCTATGACATGAGCAGACACTTCTACAGCAACGGGAGAGAGAAAGATTTACATGGAAA 100

QY 300 TTGATCTGTGACGAGAACTGAAATATTGAAAGCGGAAATGGCACTGATGAAACATTTG 359  
Db 101 TTGATCCCAATACCAAGACAGAAATATTGAAAGTGGAAATGGCACTGATGAAACATTTG 160

QY 360 AAGTGACGAGCTTTAAAGACGATACACTGCGCATCTACTCTGGGTCTTCAAAATGTT 419  
Db 161 AAGTCCATGACTTTAAAGATGAATACACTGGCATCTACTTTGTAGGTCTTCAAAATGCT 220

QY 420 TTATCAAACTCAGATTAAGTGATTTCTGTAATTTCTGACCAAGAGAGAGAAATAGATG 479  
Db 221 TTATTAACCTCAATCAAGTGATTTCTGTAATTTCTGACCAAGAGAGAGAAATAGATG 280

QY 480 AGAATGAAGAAATTAACCAACTTTCTTTGAACAGTCAGTGATTTGGTCCCGACAGAAA 539  
Db 281 AGAATGAAGAAATTAACCAACTTTCTTTGAACAGTCAGTGATTTGGTCCCGACAGAAA 340

QY 540 AGCCTATTGAAACCGAGATTTCTTAAATTTCCAAATTTCTGGAGATTTGTGATTAACG 599  
Db 341 AGCCTATTGAAACCGAGATTTCTTAAATTTCTGTAATTTCTGACCAAGAGAGAAATAG 400

QY 600 TGACCATGTTATGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCAACAGCTTTTCAGG 659  
Db 401 TGACCATGTTATGGATCAATCCCACTCTAATATGAGTTTCTGAGTTTCAACAGCTTTTCAGG 460

QY 660 AGAGGAGAGAGATCTTCTACTTTCTGCGCAACGAAAGAGAGAGAGAGAGAGAGAGAG 719  
Db 461 AGAGCGTGGAAGATCTTCTACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 520

QY 720 AGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
Db 521 AATGGGTGGTCCCTCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 580

QY 780 AAGAACTTCCAAATAAATGACTATATCTGAAATTTGAAATTTGATTTGATTTGATTTGAT 839  
Db 581 AAGAACTTCCAAATAAATGACTATATCTGAAATTTGAAATTTGATTTGATTTGATTTGAT 640

QY 840 AGAGAGTTATTTGTTATTTACTTCTGCGTCGAGGCAACCGCTATTTCGCCCGCGCTCTGTG 899  
Db 641 AGAGAGTTATTTGTTATTTACTTCTGCGTCGAGGCAACCGCTATTTCGCCCGCGCTCTGTG 700

QY 900 AACCTTTACTAGGCTACTACCATATCTGCTTACCAAGAGAGAGAGAGAGAGAGAGAGAG 959

```

Db      701 AACCTTTACTAGCTACTACCCATACCCCTACTGATACCAAGAGTTCGATCACTGTC 760
Qy      960 GTGTCATCATGCTTGTAACTGCT-GGGTGGCCGCGATGC-TGGGGAGGGTCTTAATAGG- 1016
Db      761 GTGTCATCATGCTTGTCAACTGCTGGGTGGCCGCGATGCTTGGGAAGAGTCTTAATAGGC 820
Qy      1017 AGGTTTGAGTCAAAATGCTTAAATCTGCTGGCA-ACATATATATAAATGCATGCTATTCAAT 1075
Db      821 AAGATTGAGTTCAAACGCTTAACTCTTGTGTACCATATTATTATGTCATGCTACTCCCT 880
Qy      1076 GAATTTCTGCCTAT 1089
Db      881 GGATTTCTGCCTAT 894

RESULT 12
CB588736
LOCUS   CB588736
DEFINITION AGENCOURT 12567644 NIH_MGC_136 Mus musculus cDNA clone
IMAGE:30290619 5', mRNA sequence.
CB588736
CB588736.1 GI:29506592
EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 845)
TITLE    NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL  National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT  Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM331 row: j column: 04
High quality sequence stop: 738.

FEATURES
         source
         1..845
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            /clone="IMAGE:30290619"
            /tissue type="embryonic limb, maxilla and mandible"
            /lab host="DH10B (phage-resistant)"
            /clone lib="NIH MGC 136"
            /note="Vector: PCMV-SF076.1; Site 1: EcoRV; Site 2: NotI;
            Normalized, full-length enriched library from pool of
            mouse embryonic limb, maxilla and mandible, embryonic day
            17.5, 18.5 and newborn (mandible 5, 4 and 1 limb and jaw
            equivalents from respective days). Cloned directionally,
            oligo-dT primed (5'-GACTAGTCTTAGATCGAGCGCGCGCC(7)15-3',
            size selected for the >1kb fragments, average insert size
            1.2 kb. Normalization to Cot 7.5. Tissue contributed by
            David Rowe; library constructed by Resgen, Invitrogen
            Corp. Note: this is a NIH_MGC Library."

ORIGIN

Query Match          54.4%; Score 650.6; DB 14; Length 845;
Best Local Similarity 89.0%; Pred. No. 1e-114;
Matches 712; Conservative 0; Mismatches 84; Indels 4; Gaps 1

Qy      402 TGGGCTCTTCAAAAATGTTTATCAAAACACAGATTAAAGTGAATTCCTGAATTTCTGAAAC 461
Db      1   TAGGCTCTTCAAAAATGCTTTATTATAAACTCAATCAAGTGAATTCCTGAATTTCTGAAAC 60

Qy      462 CAGAAGAGGAATAGATGAGATGAGAAATACCACAACTTCTTTGACAGTCAGTCA 521

```

Tel: 615 936 2660  
 Fax: 615 936 2661  
 Email: al.george@vanderbilt.edu  
 Insert Length: 1647 Std Error: 0.00  
 Seq primer: T3: ATTAACCTCACTAAAGGA  
 High quality sequence start: 69  
 High quality sequence stop: 741.  
 Location/Qualifiers

## FEATURES

source  
 1..803  
 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9615"  
 /clone="CH3#065\_G01"  
 /tissue\_type="heart"  
 /cell\_type="heart"  
 /dev\_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"  
 /clone\_lib="Canine heart normalized cDNA Library in pBluescript"  
 /note="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

## ORIGIN

Query Match 52.1%; Score 623.8; DB 14; Length 803;  
 Best Local Similarity 91.1%; Pred. No. 1.4e-109;  
 Matches 674; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 449 GAATTTCTGACCAAGAGGAAGAAATAGATGAGATGAAGAAATPACCACAACTTCTTT 508  
 |||||  
 Db 803 GAATTTCTGACCAAGAGGAAGAAATAGATGAGATGAAGAAATCCCACTTCTT 744  
 |||||  
 QY 509 GAACAGTCAGTATTTGGTCCACGAGAAAGCCTATTGAAACCGAGATTTCTTAAA 568  
 |||||  
 Db 743 GAACAGTCTGATTTGGTCCACGAGAAAGCCTATTGAAATCGAGACTTTCTGAAA 684  
 |||||  
 QY 569 AATTCACAAATTTCTGAGATTTGTGATAACGTGACCATGATTCGATCAATCCCACTTA 628  
 |||||  
 Db 683 AATTCACAAATTTCTGAAATTTGTGTAATGTGACCAATGATTTGATGATCAATCCCACTTA 624  
 |||||  
 QY 629 ATATCAGTTCTGAGTACAGACTTTGAGAGGAGGAGAGATCTTCACATTTCTGCC 688  
 |||||  
 Db 623 ATAGCAGTTTTCAGAGTTTACAAGACTTTTGAGAGAGATGCTGAAGATCTTCATTTCTTACC 564  
 |||||  
 QY 689 AAGCAAAAAAGGATTTGAACAAATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAG 748  
 |||||  
 Db 563 AATGAAAAAAGATTTTGAACAAACGACAGTGGTGGTCCCAAGTGAAGATGAA 504  
 |||||  
 QY 749 AAGACCCGTCAACGACACAGCAAGTGAAGAACTTCCAATAAATGACTATACCTGAA 808  
 |||||  
 Db 503 AAGATCCGACACACAGCAAGCAAGTGAAGAACTTCCAATAAATGACTATACCTGAA 444  
 |||||  
 QY 809 AATGGATAGAAATTTGATCCATCTGATGAGAGAGGTTATTGTTGTTTACTGCGGT 868  
 |||||  
 Db 443 AATGGACTAGAAATTTGATCCATCTGATGAGAGGTTATTGTTGTTTACTGCGGT 384  
 |||||  
 QY 869 CGAGGCAACCGCTATTGCGCGCGCGCTGTGAACCTTTACTAGCTACTACCAATATCCA 928  
 |||||  
 Db 383 CGAGGCAACCGCTACTGCGCGCGCGCTGTGAACCTTTACTAGCTACTACCAATATCCA 324  
 |||||  
 QY 929 TACTGCTTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCTTGTACTGCTGCGGTG 988  
 |||||  
 Db 323 TACTGCTTACCAAGGAGGCGGGTTACTGTCGTGTCATCATGCTTGTACTGCTGCGGTG 264  
 |||||  
 QY 989 GCCCGCATGCTGGGAGGGTCTTAATAGAGGTTTGAAGCTCAAAATGCTTAAA---CTGCTG 1045  
 |||||  
 Db 263 GCCCGCATGCTGGGAGGGTCTTAAGGCGCAAGTTTGAGCTCAAGTGTCTTAACTTCTGCGAG 204  
 |||||  
 QY 1046 GCAACATATAATAAATGCAATGCTATTCAATGAATTTCTGCCTATGAGGCAATCTGGGCCCT 1105  
 |||||

Db 203 CAAACATATATAAATGATGCTATCCATGAATTTCTGCCTATGAGGCAATTTGGTCTCT 144  
 QY 1106 GTAGCCAGCTCTCCAGAAATTAATTTGTTAGGTAATTCCTCTCTTCAATGTTCTATAAATTT 1165  
 |||||  
 Db 143 AGTAGCCAGTACTCTCCAGAAATTAATTTGTTAGGTAATTCCTCTCTTCAATGTTCTATAAATTT 84  
 |||||  
 QY 1166 CTACATTAATATCAACCAAAAAA 1185  
 |||||  
 Db 83 CTACATCAATCAAAAAAAGAA 64

## RESULT 14

BY714257

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY714257 932 bp mRNA linear EST 17-DEC-2002  
 BY714257 RIKEN full-length enriched, 0 day neonate head Mus  
 musculus cDNA clone 4833424007 5', mRNA sequence.

BY714257.1 GI:27127366

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (Bases 1 to 932)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,  
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,  
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,  
 Clothier, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,  
 Fletcher, C. P., Forrest, A., Frazer, K. S., Gaasterland, T. A.,  
 Gariboldi, M., Gliss, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,  
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,  
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,  
 Petrovsky, N., Fillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,  
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,  
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,  
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
 Wells, C., Wilm, L. G., Wyszewski, B., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
 Hayashizaki, Y., Hironaka, T., Konno, H., Nakamura, M.,  
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,  
 Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 MEDLINE  
 PUBMED  
 COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Hirozane, T., Hori, F.,  
 Imotani, K., Hashizume, W., Hayashida, K., Kawai, J., Kojima, Y.,  
 Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
 Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,  
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
 Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

```
FEATURES
source
1..932
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="14833424007"
  /sex="mixed"
  /tissue_type="head"
  /dev_stage="0 day neonate"
  /lab_host="DH10B"
  /clone_lib="RIKEN full-length enriched, 0 day neonate head"
  /note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCTCGGTTAATAAATTAATCCGCCCCCC sequence [5' GAGAGAGATCTCGGTTAATAAATTAATCCGCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN
Query Match 52.08; Score 623; DB 13; Length 932;
Best Local Similarity 83.94; Pred. No. 1.9e-109;
Matches 726; Conservative 0; Mismatches 137; Indels 2; Gaps 2;

QY 25 AAAGCAGGAAAGAGTACTGTGCTCAGAGACCATGGCAAGAAATCCTCCAGAAATTG 84
DB 1 AAAGCTGGAAAGAGCACCGGTGCTGGAGACCATGGCAAGATTCCTCCAGAACTG 60
QY 85 TGAAGATGTACATTTCTAAATCAGAAAGCTTTTAAATCCAGAAATATGTAAATCACT 144
DB 61 TGAGGGCTGTACATTTCTAAATCAGAAAGCTCTGATATCTAAGATGATATGTTAATTA 120
QY 145 TAGATTTTGGCATGCTGTTGTATCTTCGGCCCTAACTTAATTCCTGTTTTCGGG 204
DB 121 GAAGATTGTGGATAGTGGTTGGTATCCTGGCCCTTAACCTAATTTGTTCTGTTTGGG 180
QY 205 GAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCCTATGACATGGAGCACTTTCTA 264
DB 181 GAGCAACACTTCTGGCCCGCATGTATCCAGATAACCTATCATATGGAGCACTTTCTA 240
QY 265 CAGCAATGGAGAGAGAGAGATTACATGGAATTTGATCTGTGACCAGAACTGAAAT 324
DB 241 CAGCAACGGCAGAGAGAGAGATTACATGGAATTTGATCTGTGACCAGAACTGAAAT 300
```

```
QY 325 ATTCAGAGCGGAATGGCACTGATGAAACATTTGGAAGTGCAGCACTTTAAAAACGGATA 384
DB 301 ATTGAGAGTGGAAATGGCACTGATGAAACATTTGGAAGTGCAGCACTTTAAAAATGGATA 360
QY 385 CACTGGCATCTACTTCCTGGGTCTTCAAAAATGTTTATCAAA-ACTCAGATTAAAGTGA 443
DB 361 CACTGGCATCTACTTCCTGGGTCTTCAAAAATGTTTATTAATAACTCAATCAAAAGTGA 420
QY 444 TTCTGTAATTTCTGAAACAGAGAGAAATAGATGAGAATGAAGAAATTTACCACAATTT 503
DB 421 TTCTGTAATTTCTGAAACAGAGAGAAATAGATGAGAATGAAGAAATTTACTACAATTT 480
QY 504 TCTTTGAAACAGTCAGTCAGTTTGGGTCCCGACGAGAAAAGCCTATTGAAAACCGAGATTTTC 563
DB 481 TTTTGTGAACAGTCAGTCAGTTTGGGTCCCGACGAGAAAAGCCTATTGATACAGAGACTTCC 540
QY 564 TTAATAATTTCAAATAATTTCTGGAGATTTGTAAGCTGACCATGATTTCGATCAATCCCA 623
DB 541 TGAATAATTTCTAATATTTCTGGAGATTTGCGAATAATGTGACCATGATTCTGATCAATCCCA 600
QY 624 CTCTAATATCACTTTCTGAGTTTACAAGACTTTTGAGGAGGAGGAGAGATCTTCACTTTC 683
DB 601 CTCTAATATCACTTTCTGAGTTTACAAGACTTTTAGGAGGAGGAGAGATTTCTCACTCTC 660
QY 684 CTGCCAACGAAAAAAGGGATTGAACAAAATGAACAGTGGGTGCTCCCTCAAGTGAAG 743
DB 661 CTACCAAGTGAATAAAGGGATTGAACAAAAGCAATGGGTGGTCCGCAAGTGAAGG 720
QY 744 TAGAGAGACCGGTCTACGCGACAGACAGCAAGTGAAGAACTTCCAAATAATGACTATA 803
DB 721 TGGAGAAG-CCCGCCACACACAGCAAGCAAGCAAGCAAGTGGGTGGTCCGCAAGTGAAG 779
QY 804 CTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGGTTATTTGTTGTTACT 863
DB 780 CTGAATGAGAGTGGATATGACCATAGCCGAGAGAAATGCAAGTGTGTTTCTCTGTCG 839
QY 864 GCGTCTGAGGCAACCGCTATTGCGG 888
DB 840 CGAGGCAACGTTACGCGCGTGGCG 864

RESULT 15
CB587595 876 bp mRNA linear EST 03-APR-2003
LOCUS AGENCOURT 12771379 NIH MGC 136 Mus musculus cDNA clone
DEFINITION IMAGE:30295278 5', mRNA sequence.
CB587595
VERSION CB587595.1 GI:29505451
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. (bases 1 to 876)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM343 row: 1 column: 07
High quality sequence stop: 660.
Location/Qualifiers
1..876
/organism="Mus musculus"
/mol_type="mRNA"
```

/db\_xref="taxon:10090"  
/clone="IMAGE:30295278"  
/tissue type="embryonic limb, maxilla and mandible"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_136"  
/note="Vector: PCWV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;  
Normalized, full-length enriched library from pool of  
mouse embryonic limb, maxilla and mandible, embryonic day  
17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw  
equivalents from respective days). Cloned directionally,  
oligo-dT primed (5'-GACTAGTCTAGATCGGCGGCCCTT)15-3'.  
Size selected for the >1kb fragments, average insert size  
1.2 kb. Normalization to Cot 7.5. Tissue contributed by  
David Rowe; library constructed by ResGen, Invitrogen  
Corp. Note: this is a NIH\_MGC Library."

ORIGIN	Query Match	Score	620.6;	DB 14;	Length	876;
	Best Local Similarity	89.4%;	Pred. No. 5.7e-109;			
	Matches	680;	Conservative	0;	Mismatches	79;
					Indels	2;
					Gaps	1;
QY	11	CTCTAGTCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACATGGCAAGAAT	70			
Db	16	CTCTAGTCTCTCAAGCAAGGAAGAGACCGTGTGCTGGAGACCATGGCAAGAAT	75			
QY	71	CTCCAGAGAAATGTGAAGACTGTCAATCTAAATCAGAAGCTTTTAAATCCAAGAAA	130			
Db	76	CTCCAGAGAACTGTGAGGGCTGTCAATCTAAATCAGAAGCTCTGAAATCTAAGAAG	135			
QY	131	ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCTCTGGCCCTAACTTAATT	190			
Db	136	ATATGTAATCACTTAAGATTGTGGACTAGTGTGGTATCTCTGGCCCTTAATCTTAATT	195			
QY	191	GTCTCTTTTGGGGAGCAAGCACTTCTGCGCCGAGGTACCCAAAAGCCTATGACATG	250			
Db	196	GTCTCTTTTGGGGAGCAACACTTCTGCGCCGAGGTATCCAGAAAACCTATGACATG	255			
QY	251	GAGCACACTTCTACAGCAATGAGAGAGAAAGATTTTACATGGAATTTGATCTCTGTG	310			
Db	256	GAGCACACTTCTACAGCAACGGCGAGAGAAAGATTTTACATGGAATTTGATCCATA	315			
QY	311	ACCAGAACTGAATATTCAGACGCGAATGGCACTGATGAACATTTGGAAGTGCAGAC	370			
Db	316	ACCAGAACTGAATATTCAGAGTGAATTTGCAAGTGAATGGCACTGATGAACATTTGGAAGTGCATGAC	375			
QY	371	TTTAAAAACGATACACTGGCATCTACTCTGCTGGGTCTTCAAAAATGTTTATCAAAACT	430			
Db	376	TTTAAAAATGGATACACTGGCATCTACTTTGTAGGTCTTCAAAAATGCTTTATTAATACT	435			
QY	431	CAGATTAAAGTATCTCGAATTTCTGAACCGAAGAGAAATATAGATGAGATGAAGAA	490			
Db	436	CAAAATCAAGTATCTCGAATTTCTGAACCGAAGAGAAATATAGATGAGATGAAGAA	495			
QY	491	ATTACCAACACTTCTTTCACAGCTCAGTGATTGCGTCCCGACAGAAAAGCCTATTGAA	550			
Db	496	ATTACTACAACTTCTTTCGACAGCTCAGTGATTGCGGTCCCGACAGAAAAGCCTATTGAA	555			
QY	551	AACCGAGATTTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATAACCTTGACCATGTAT	610			
Db	556	AACAGAGACTTCTGAAAAATTTCTAAAAATTTCTGGAGATTTGCGATAATGTGACCATGTATC	615			
QY	611	TGGATCAATCCCACTCTAATATACAGTTTCTGAGTTTACAAAGCTTTTGGAGGAGGGAGAA	670			
Db	616	TGGATCAATCCCACTCTAATAGCAGTTTACAGGAATTTACAGGACTTTTGGAGGAGCGGTGAA	675			
QY	671	GATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAACAAAATGAACAGTGGGTGG--	728			
Db	676	GATCTTCACTTTCTACCACTGAAAAAAGGATTTGACCAAGATGAGCAATGGGGGGT	735			
QY	729	TCCCTCAAGTGAAGTAGAGAGACCCCGTCAGCCAGACAA	769			
Db	736	CCCGCCAAAGTGAAGTGAAGAGACCCCGCCACACAGAA	776			

RESULT 16  
CD512100  
LOCUS  
DEFINITION  
CD512100.1  
AGENCOURT 14354032 NIH MGC 186 Homo sapiens cDNA clone  
IMAGE:30406681 5', mRNA sequence.  
CD512100  
EST.  
CD512100.1  
GI:31443818  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 751)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: gcapos-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCM195 row: n column: 02  
High quality sequence stop: 553.  
Location/Qualifiers  
1. .751  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30406681"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfII  
(ggccatgatgccc); Site 2: SfII (ggccgctcggcc); Library is  
oligo-dT primed and directionally cloned. cDNA was  
prepared from a pooled samples of tissues from Skin,  
meninges, duramater, pia mater and choroid plexus.  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library"

ORIGIN	Query Match	Score	612.6;	DB 14;	Length	751;
	Best Local Similarity	99.2%;	Pred. No. 2e-107;			
	Matches	626;	Conservative	0;	Mismatches	4;
					Indels	1;
					Gaps	1;
QY	237	AAGCCATATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGATTACATCG	296			
Db	54	AGGCCATATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGATTACATCG	113			
QY	297	AAATTGATCTCTGTGACCAAGAACTGAATATTTCAGAGCGGAATGGCACTGATGAACAT	356			
Db	114	AAATTGATCTCTGTGACCAAGAACTGAATATTTCAGAGCGGAATGGCACTGATGAACAT	173			
QY	357	TGGAAGTACAGACTTTTAAAAACGATACACTGGCATCTTCTCGTGGGTCTTCAAAAT	416			
Db	174	TGGAAGTACAGACTTTTAAAAACGATACACTGGCATCTTCTCGTGGGTCTTCAAAAT	233			
QY	417	GTTTTATCAAAACTCAGATTAAAGTGATTCTCTGAATTTTCTGAACCAAGAGGAATAG	476			

```
Db 234 GTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACACAGAGGAAATAG 293
QY 477 ATGAGATGAAGAAATACCAACATTTCTTTGAACAGTCAGTGATTTGGTCCCGACAG 536
Db 294 ATGAGAAATGAAGAAATACCAACATTTCTTTGAACAGTCAGTGATTTGGGTCCCGACAG 353
QY 537 AAAGCCCTATTGAACACCGAGATTTCTTAAAAATTCMAAATTCGAGATTTGTGATA 596
Db 354 AAAAGCCCTATTGAACACCGAGATTTCTTAAAAATTCMAAATTCGAGATTTGTGATA 413
QY 597 ACCTGACCATGTATTGGATCAATCCACTCTAAATATCAATTCAGTTTCTGAGTTCAAGACTTTG 656
Db 414 ACCTGACCATGTATTGGATCAATCCACTCTAAATATCAATTCAGTTTCTGAGTTCAAGACTTTG 473
QY 657 AGGAGGAGGAGAGATCTTCACCTTCTGCAACGAAAAAAGGATTCGAACAAATG 716
Db 474 AGGAGGAGGAGAGATCTTCACCTTCTGCAACGAAAAAAGGATTCGAACAAATG 533
QY 717 AACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAAGCAAGTG 776
Db 534 AACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAAGCAAGTG 593
QY 777 AGGAAGAACTTCCATAAATGACTATACCTGAATATGGAATGGAATTCATCCCATGCTGG 836
Db 594 AGGAAGAACTTCCATAAATGACTATACCTGAATATGGAATGGAATTCATCCCATGCTGG 653
QY 837 ATGAGAG-AGTTTATGTTGTTATTACTGCC 866
Db 654 ATGAGAGAAATTTATGGTTGTTATTACTGCC 684

RESULT 17
CB587917 867 bp mRNA linear EST 03-APR-2003
LOCUS AGENCOURT_12990010 NIH_MGC_136 Mus musculus cDNA clone
DEFINITION IMAGE:30292458 5', mRNA sequence.
ACCESSION CB587917
VERSION CB587917.1 GI:29505773
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 867)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Cloning: Agencourt Bioscience Corporation
Cloning Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDAM336 row: f column: 19
High quality sequence stop: 605.
Location/Qualifiers
1..867
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30292458"
/tissue_type="embryonic limb, maxilla and mandible"
/lab_host="NIH MGC 136"
/clone_lib="NIH MGC 136"
/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized, full-length enriched library from pool of
mouse embryonic limb, maxilla and mandible, embryonic day
17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw
equivalents from respective days). Cloned directionally,
oligo-dT primed (5'-GACTAGTTCTAGATCGGCGCGCCGCC (7)15-3'.
```

```
Size selected for the >1kb fragments, average insert size
1.2 kb. Normalization to Cot 7.5. Tissue contributed by
David Rowe; library constructed by Resgen, Invitrogen
Corp. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 50.8%; Score 608.2; DB 14; Length 867;
Best Local Similarity 91.1%; Pred. No. 1.3e-106; Indels 0; Gaps 0;
Matches 646; Conservative 0; Mismatches 63;

QY 47 TGCTGAGAGACCATGGCAAGAAATCCCTCCAGAGAAATGTGAAGACTGTCAACATCTCTAAAT 106
Db 1 TGCTGGGAGACCATGGCAAAATCTCTCCAGAGAACTGTGAGGGCTGTCAACATCTCTAAAT 60
QY 107 GCAGAAAGCTTTAAATCCAGAAAAATATGTAATCACTTAAGATTTGTGGAGCTGTGTTT 166
Db 61 GCAGAAAGCTCTGAAATCTTAAGAGATATGTAATCACTGAAAGATTTGTGGAGCTGTGTTT 120
QY 167 GGTATCTCTGGCCCTAACTCTAAATTTGTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAG 226
Db 121 GGTATCTCTGGCCCTTAACCTCTAAATTTGTCTGTTTGGGGGAGCAAACTTCTGGCCGAG 180
QY 227 GTACCCAAAAAGCCTATGACATGAGCAGACACTTTCTACAGCAATGAGAGAAAGAAAG 286
Db 191 GTATCCAAAGAAACCTATGATGAGCAGCACTTTCTACAGCAACGCGGAGAGAAAGAAAG 240
QY 287 ATTTACATGGAATTTGATCTCTGTCACGAGCACTTTCTACAGCAATGAGAGAAAGAAAG 346
Db 241 ATTTACATGGAATTTGATCTCTGTCACGAGCACTTTCTACAGCAACGCGGAGAGAAAGAAAG 300
QY 347 GATGAACATTTGGAAGTGCAGCACTTTTAAACCGATACACTGGCATCTACTTCTGGGT 406
Db 301 GATGAACATTTGGAAGTGCAGCACTTTTAAATGATGATGATGATGATGATGATGATGATGAT 360
QY 407 CTTCAAAATGTTTATCAAACTCAGATTTAAAGTGAATTCCTGAAATTTCTGAACAGAA 466
Db 361 CTTCAAAATGCTTTTATTAACCTCAATCAAGTGAATTCCTGAAATTTCTGAACAGAG 420
QY 467 GAGGAAATAGATGAGAAATTAACCAACTTTCTTTGAAACGATGATGATGATGATGATGATGAT 526
Db 421 GAAGAAATAGATGAGAAATTAACCAACTTTCTTTGAAACGATGATGATGATGATGATGATGAT 480
QY 527 GTCCAGCAGAAAGCCCTATTGAACCGAGATTTCTTAAATTTCCAAAATTTCTGGAG 586
Db 481 GTTCCCGCAGAAAGCCCTATTGAACCGAGACTTCTCTGAAAAATTTCTAAAAATTTCTGGAG 540
QY 587 ATTTGTGATAACGTGACCATGATTTGGATCAATCCCACTCTCTAATATCACTTTCTGAGTTA 646
Db 541 ATTTGGGATTAATGTGACCATGATTTGGATCAATCCCACTCTCTAATATCACTTTCTGAGTTA 600
QY 647 CAAGACTTTGAGAGAGGAGAGAGATCTTCACTTTCTCTGCCAACGAAAAAAGGAGTT 706
Db 601 CAGGACTTTGAGAGAGAGAGAGATCTTCACTTTCTCTGCCAACGAAAAAAGGAGTT 660
QY 707 GAACAAATGAAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAGAG 755
Db 661 GACCAGAAATGAACATGAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAG 709

RESULT 18
BI454468 744 bp mRNA linear EST 21-AUG-2001
LOCUS 603171428F1 NCI_OGAP_Mam5 Mus musculus cDNA clone IMAGE:5250871 5',
DEFINITION mRNA sequence.
ACCESSION BI454468
VERSION BI454468.1 GI:15245124
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11633 row: h column: 08  
High quality sequence start: 2  
High quality sequence stop: 733.  
High quality sequence stop: 733.

FEATURES  
source  
Location/Qualifiers  
1. 744  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5250871"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam5"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

ORIGIN  
Query Match 48.4%; Score 579.6; DB 12; Length 744;  
Best Local Similarity 90.7%; Pred. No. 4.2e-101;  
Matches 640; Conservative 0; Mismatches 64; Indels 2; Gaps 2;

QY 11 CTCTCAGTCTCTCAAGACGAGGAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 70  
DB 35 CTCTCAGTCTCTCAAGACGAGGAAGAGACCGTGTGTGGGAGACCATGGCAAGAAT 94  
QY 71 CCTCCAGAGAAATGTGAAGACTGTCAATCTAAATGCAGAACTTTTAATCCAGAAA 130  
DB 95 CCTCCAGAGAACTGTGAGGGCTGTCAATCTAAATGCAGAACTCTGAATCTAAGAAG 154  
QY 131 ATATGTAATCACTTAAGAATTTGTGACTGGTGTGGTATCTTGGCCCTAACTCTAATT 190  
DB 155 ATATGTAATCACTTAAGAATTTGTGACTAGTGTGGTATCTTGGCCCTAACTCTAATT 214  
QY 191 GTCTCTGTTTGGGGAGCAGACACTTCTGCCCGAGGTACCCAAAAGCCTATGACATG 250  
DB 215 GTCTCTGTTTGGGGAGCAACACTTCTGCCCGAGGTATCCAAAGAAAACCTATGACATG 274  
QY 251 GAGCACACTTCTACAGCAATGAGAGAGAAAGAAATTTACATGGAATTTGATCTGTG 310  
DB 275 GAGCACACTTCTACAGCAACGCGAGAGAGAAAGATTTACATGGAATTTGATCTGTG 334  
QY 311 ACAGAACTGAAATATTCAGACGGAATGGCACTGATGAACATTTGGAAGTGCACGAC 370  
DB 335 ACCAGAAACGAAATATTCAGAAAGTGAATGGCACTGATGAACATTTGGAAGTGCATG 394  
QY 371 TTTAAAACGGAATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 430  
DB 395 TTTAAAACGGAATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 454  
QY 431 CAGATTAAAGTGAATCTCTGAATTTCTGAACACGAGAGAGAAATAGATGAGATGAGAA 490  
DB 455 CAAATCAAGTGAATCTCTGAATTTCTGAACACGAGAGAGAAATAGATGAGATGAGAA 514  
QY 491 ATTACCAACTTCTTTGAACAGTCAAGTGAATTTGGTCCCGAGCAAAAAGCCTATTGAA 550  
DB 515 ATTACTACAACCTTCTTTGAACAGTCAAGTGAATTTGGTCCCGAGCAAAAAGCCTATTGAA 574  
QY 551 AACCGAGATTTCTTAAAATTCMAAATCTGGAGATTGTG-TGATAACGTGACCATGTA 609

DB 575 AACAGAGACTTCTGAAATAATCTAAATTTCTGGAGATTTGCCGATATGTGACCATGTA 634  
QY 610 TTGATCAATCCCACTCTPAA-TATCAGTTTCTGAGTTTACAGACTTTGAGGAGGAGGAG 668  
DB 635 CTGATCAATCCCACTCTTAATAGCAGTTTCAGAAATTACAGACTTTTGGAGGAGCGGTG 694  
QY 669 AAGATCTTCACTTCTCCGCCCAACGAAAGAAAAGGATGAACAAA 714  
DB 695 AAGATCTTCACTTCTTCTACCAGTGAACAAAAGGGGATGACCCAGA 740

RESULT 19  
BF182340  
LOCUS BF182340  
DEFINITION 601804094F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:4035214 5',  
mRNA sequence.  
BF182340  
VERSION BF182340.1 GI:11060482  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (Bases 1 to 846)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9308 row: k column: 23  
High quality sequence stop: 703.  
Location/Qualifiers  
1. 846  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4035214"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam5"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

FEATURES  
source

Location/Qualifiers  
1. 846  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4035214"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam5"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

## ORIGIN

Query Match 48.4%; Score 579; DB 10; Length 846;  
Best Local Similarity 88.2%; Pred. No. 5.3e-101;  
Matches 675; Conservative 0; Mismatches 85; Indels 5; Gaps 4;

QY 11 CTCTCAGTCTCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 70  
DB 32 CTCTCAGTCTCTCAAGCAAGAAAGAGACCGTGTGTGGGAGACCATGGCAAGAAT 91  
QY 71 CCTCCAGAGAAATGTGAAGACTGTCAATTTCTAAATGCAGAGCTTTAAATCCAGAAA 130  
DB 92 CCTCCAGAGAACTGTGAGGGCTGTCAATTTCTAAATGCAGAGCTCTGAAATCTAAGAAG 151  
QY 131 ATATGTAATCACTTAAGAATTTGTGACTGGTGTGGTATCTTGGCCCTAACTCTAATT 190  
DB 152 ATATGTAATCACTTAAGAATTTGTGACTAGTGTGGTATCTTGGCCCTAACTCTAATT 211



```
QY 191 GTCTGTTTGGGGAGCAACACTTCTGGCCGAGGTACCCAAAAAGCCTATGACATG 250
Db 212 GTCTGTTTGGGGAGCAACACTTCTGGCCGAGGTATCCAGAAACCTATGACATG 271
QY 251 GAGCACACTTCTACAGCAATGAGAGAGAGAGATTATCATGGAAATTTGATCTCTGTG 310
Db 272 GAGCACACTTCTACAGCAACGGCGAGAGAGAGATTATCATGGAAATTTGATCTCTGTG 331
QY 311 ACCGAACTGAAATATTCAGAGCGGAATGGCACTGATGAAATTCGGAAGTGCACGAC 370
Db 332 ACCGAACTGAAATATTCAGAGCGGAATGGCACTGATGAAATTCGGAAGTGCACGAC 391
QY 371 TTTAAAAACGGATACACTGGCATCTACTTCTGGGTCTTTCAAAAATGTTTATCAAAACT 430
Db 392 TTTAAAAATGATACACTGGCATCTACTTCTGGGTCTTTCAAAAATGTTTATCAAAACT 451
QY 431 CAGATTAAAGTGAATTCCTGAAATTTCTGAAACGAGAGAGAAATAGATGAGAAATGAAGA 490
Db 452 CAAATCAAAAGTGAATTCCTGAAATTTCTGAAACGAGAGAGAAATAGATGAGAAATGAAGA 511
QY 491 ATTACCAACAATTTCTTTGAAACAGTCAGTGATTTGGGTCCCGAGAGAAAGCCTATTGAA 550
Db 512 ATTACCAACA--TTTCTTTGAAACAGTCAGTGATTTGGGTCCCGAGAGAAAGCCTATTGAA 570
QY 551 AACCAGAGATTTTCTTAAAAATTCCTGAAATTTCTGAGATTTGGAT--AACGTGACCATGTA 609
Db 571 AACAGAGACTTCTGAAATTTCTTAAAAATTTCTGAGATTTGGATTAATGTGACCATGTA 630
QY 610 TTGGATCAATCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGAGAGGGAGA 669
Db 631 CTGGATCAATCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGAGAGGGAGA 690
QY 670 AGATCTTCACTTCTCTGCCAACGAAAAAAGGATTTGAAACAAATTAAGACAGTGGGTGGT 729
Db 691 AGATCTTCACTTCTTACACAGTG--AAAAAGGGGATTTGACAGATTCAGCATGCGGGGGT 749
QY 730 CCTCAAGTGAAGTAGAAGAACCGCTCAGCCGACGACAGCAAGCAAG 774
Db 750 CC--GCAGTGGGGGGGGGAAAAACCGGCAACAGAAAGAGGAGG 792

RESULT 20
BG174435
LOCUS 60233425F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457652 5',
DEFINITION mRNA sequence.
ACCESSION BG174435
VERSION BG174435.1 GI:12681138
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 958)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10254 row: m column: 13
High quality sequence start: 10
High quality sequence stop: 652.
Location/Qualifiers
1..958
/organism="Mus musculus"
/mol_type="mRNA"
```

```
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457652"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match 48.2%; Score 577; DB 12; Length 958;
Best Local Similarity 89.3%; Pred. No. 1.2e-100;
Matches 633; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 41 ACTGTGCTCAGAGACCATGCAAGAAATCCTCCAGAGAAATTTGTGAAGACTGTGCACATT 100
Db 2 ACCGTGTCTGGGAGACCATGCAAGAAATCCTCCAGAGAACTGTGAGGCTGTGCACATT 61
QY 101 CTAATGAGAGAGCTTTTAAATCCAGAAATAATGTAATCACTTAAGATTGTGGACTG 160
Db 62 CTAATGAGAGAGCTCTGAATCTAAGAGATATGTAATCACTGAAGATTGTGGACTA 121
QY 161 GTGTTGTTGATTCCTGGCCCTAATCTAATTTGCTGTTTGGGGGAGCAAGCACTTCTGG 220
Db 122 GTG--TTGGTATCCTGGCCCTTAACTCTAATTTGCTGTTTGGGGGAGCAAGCACTTCTGG 180
QY 221 CCGGAGGTACCCAAAAAGCCCTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAG 280
Db 181 CCGGAGGTATCCAGAAACCTATGACATGGAGCACACTTTCTACAGCAACGCGGAGAAG 240
QY 281 AAGAAGATTTACATGGAATTTGATCCTGTGACAGAACTGAAATATTGAGAGCGGAAT 340
Db 241 AAGAAGATTTACATGGAATTTGATCCTGATACAGAACTGAAATATTGAGAGCGGAAT 300
QY 341 GGCACGTGAAACATTTGGAAGTGCACGACTTTTAAACCGGATACACTGGCATCTACTTC 400
Db 301 GGCACGTGAAACATTTGGAAGTGCACGACTTTTAAACATGGATACACTGGCATCTACTTC 360
QY 401 GTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGAATTCCTGAATTTCTGAA 460
Db 361 GTAGGTCTTCAAAAATGCTTTTATTAACCTCAAACTCAAACTGATTCCTGAATTTCTGAA 420
QY 461 CCAGAGAGGAAATAGATGAGAAATGAAGAAATTAACAACTTTCTTTGAAAGTCAGTG 520
Db 421 CCAGAGGAAAGAAATAGATGAGAAATGAAGAAATTAACAACTTTCTTTGAAAGTCAGTG 480
QY 521 ATTGGGTCCAGCAGAAAGCCCTATTGAAACCGGAGATTTCTTAAATAATTCCTCAAAAT 580
Db 481 ATTGGGTCCCGCAGAAAGCCCTATTGAAACCGGAGACTTCTGAAATAATTCCTCAAAAT 540
QY 581 CTGGAGATTGTGATAACGTCACCATGATTTGGATCAATCCCACTCTTAATATCAGTTTCT 640
Db 541 CTGGAGATTGTGATAACGTCACCATGATTTGGATCAATCCCACTCTTAATATCAGTTTCT 600
QY 641 GAGTTTCAAGACTTTGAGGAGGAGGAGAGATCTTCACTTTCTGCAACGAGAAAGAAA 700
Db 601 GAATTCAGGACTTTGAGGAGGAGGAGAGATCTTCACTTTCTGCAACGAGAAAGAAA 660
QY 701 GGGATTGAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGA 749
Db 661 GGGATTGGCCAGAAATGAGCAATGGTGTCCCGCAGCTGAGGTGAGAA 709

RESULT 21
CD103859 712 bp mRNA linear EST 15-MAY-2003
LOCUS AGENCOURT_14008022 NIH_MGC_186 Homo sapiens cDNA clone
DEFINITION IMAGE:30372251 5', mRNA sequence.
ACCESSION CD103859
VERSION CD103859.1 GI:30757033
```



```

KEYWORDS      EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens

REFERENCE      1 (bases 1 to 1712)
AUTHORS        Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL        NIH-MGC http://mgs.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapsb@mail.nih.gov
               Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
               cDNA Library Preparation: CLONTECH Laboratories, Inc.
               DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
               Clone Distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: NDCM164 row: c column: 12
               High quality sequence stop: 494.
               Location/Qualifiers
                 1..712
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:30372251"
                   /lab_host="DH10B (T1 phage-resistant)"
                   /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: Sfil
                   (ggccattatggcc); Site 2: Sfil (ggccctcgccg); Library is
                   oligo-dT primed and directionally cloned. cDNA was
                   prepared from a pooled samples of tissues from skin,
                   meninges, duramater, pia matter and choroid plexus. 5'
                   and 3' adaptors were used in cloning as follows: 5'
                   adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor
                   sequence: 5'-ATTCTAGAGCGCGGGCGGCATG-dT(30)BN-3'
                   (where B = A, C, or G and N = A, C, G, or T). Average
                   insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
                   contained inserts by PCR. This library was enriched for
                   full-length clones and was constructed by Clontech
                   Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                   Library"

FEATURES             location/qualifiers
source              1..712
Query Match         48.0%; Score 574; DB 14; Length 712;
Best Local Similarity 98.8%; Pred. No. 5.1e-100;
Matches 589; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 237 AAGCCTATGACATGGAGCACACTTCTACAGCAATGGAGAGAGAAATTCATG 296
DB 58 AGCCCTATGACATGGAGCACACTTCTACAGCAATGGAGAGAGAAATTCATG 117
QY 297 AAATTGATCTGTGACAGAACTGAATATTTACAGAGCGGAATGGCACTGATCAACAT 356
DB 118 AAATTGATCTGTGACAGAACTGAATATTTACAGAGCGGAATGGCACTGATCAACAT 177
QY 357 TCGAAGTGCAGCACTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAAT 416
DB 178 TCGAAGTGCAGCACTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAAT 237
QY 417 GTTTTATCAAACTCAGATTAAAGTATTCCTGAATTTCTGAACAGAGAGAAATAG 476
DB 238 GTTTTATCAAACTCAGATTAAAGTATTCCTGAATTTCTGAACAGAGAGAAATAG 297
QY 477 ATGAGAAATGAAGAAATTAACACAACTTCTTTCAGACAGTCAGTGATTTGGTCCGAGCAG 536
DB 298 ATGAGAAATGAAGAAATTAACACAACTTCTTTCAGACAGTCAGTGATTTGGTCCGAGCAG 357
QY 537 AAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCAAAATTTCTGGAGATTTTGATA 596
DB 358 AAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCAAAATTTCTGGAGATTTTGATA 417
QY 597 ACGTGACCATGATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTG 656

```

```

DB 418 AGTGACCAATGATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTG
QY 657 AGGAGGAGGAGAGAGATCTTCACTTTCTCGCAACGAAAAAAGGATTTGAACAAATG 716
DB 478 AGGAGGAGGAGAGAGATCTTCACTTTCTCGCAACGAAAAAAGGATTTGAACAAATG 537
QY 717 ACAGTGGGTGCTCCTCAAGTGAAGTAGAAGACCCGTCACGCCACAGCAAGCAAGTG 776
DB 538 AACAGTGGGTGCTCCTCAAGTGAAGTAGAAGACCCGTCACGCCACAGCAAGCAAGTG 597
QY 777 AGGAAGAACTTCCAATAATGACTACTGAAAA--TGAATAGATTTGATCCCA 830
DB 598 ATGAAGAACTTCCAATAATGACTACTGAAAAATGGAATAGATTTGATCCCA 653

RESULT 22
BI248748
LOCUS      679 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION Mus musculus cDNA clone IMAGE:5148942 5',
ACCESSION  BI248748
VERSION     BI248748.1 GI:14795446
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11367 row: m column: 03
High quality sequence stop: 677.
Location/Qualifiers
source      1..679
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="IMAGE:5148942"
            /tissue_type="tumor, gross tissue"
            /dev_stage="7 months"
            /lab_host="DH10B"
            /clone_lib="NCI CGAP Mam5"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
            Library constructed by Life Technologies. Investigators
            providing samples: Lothar Hennighausen/Robin Humphreys,
            NIH"

FEATURES             location/qualifiers
source              1..679
Query Match         46.2%; Score 553.2; DB 12; Length 679;
Best Local Similarity 91.7%; Pred. No. 5e-96;
Matches 607; Conservative 0; Mismatches 53; Indels 2; Gaps 2;

QY 11 CTCTCAGTCTCTCAAAGCAAGAGAGTACTGTGCTGAGAGACCATGCAAGAAT 70
DB 19 CTCTCAGTCTCTCAAAGC-AGGAAGAGACACCGTGTCTGGGAGACCATGCAAGAAT 77
QY 71 CTCCAGAGAATTTGCAAGACTGTGCACATTCTAAATGCAAGAGCTTTTAAATCCAGAA 130
DB 78 CTCCAGAGAATCTGAGGGCTGTGCACATTCTAAATGCAAGAGCTCTGAAATCTAAGAAG 137

```

```

131 ATATGTAATCAGTAACTGATGCTGTTGGTATCTCTGGCCCTAACTCTAATT 190
132 |||||||
133 |||||||
134 |||||||
135 |||||||
136 |||||||
137 |||||||
138 ATATGTAATCAGTAACTGATGCTGTTGGTATCTCTGGCCCTAACTCTAATT 197
139 |||||||
140 |||||||
141 |||||||
142 |||||||
143 |||||||
144 |||||||
145 |||||||
146 |||||||
147 |||||||
148 |||||||
149 |||||||
150 |||||||
151 GTCCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAGGCTTATGACATG 250
152 |||||||
153 |||||||
154 |||||||
155 |||||||
156 |||||||
157 |||||||
158 GTCCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTATCCAGAAACCTATGACATG 256
159 |||||||
160 |||||||
161 |||||||
162 |||||||
163 |||||||
164 |||||||
165 |||||||
166 |||||||
167 |||||||
168 |||||||
169 |||||||
170 |||||||
171 GAGCACACTTCTCAGCAATCGGAGAGAGAGATTATCATGGAATTCATGCTCTG 310
172 |||||||
173 |||||||
174 |||||||
175 |||||||
176 |||||||
177 |||||||
178 GAGCACACTTCTCAGCAATCGGAGAGAGAGATTATCATGGAATTCATGCTCTG 316
179 |||||||
180 |||||||
181 |||||||
182 |||||||
183 |||||||
184 |||||||
185 |||||||
186 |||||||
187 |||||||
188 |||||||
189 |||||||
190 |||||||
191 ACCGAACTGAAATTCAGAGCGGAATGSCACTGATGAAACATTCGGAAGTCAGCAG 370
192 |||||||
193 |||||||
194 |||||||
195 |||||||
196 |||||||
197 |||||||
198 ACCGAACTGAAATTCAGAGCGGAATGSCACTGATGAAACATTCGGAAGTCAGCAG 376
199 |||||||
200 |||||||
201 |||||||
202 |||||||
203 |||||||
204 |||||||
205 |||||||
206 |||||||
207 |||||||
208 |||||||
209 |||||||
210 |||||||
211 TTTAAAAACGGATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 430
212 |||||||
213 |||||||
214 |||||||
215 |||||||
216 |||||||
217 |||||||
218 TTTAAAAACGGATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 436
219 |||||||
220 |||||||
221 |||||||
222 |||||||
223 |||||||
224 |||||||
225 |||||||
226 |||||||
227 |||||||
228 |||||||
229 |||||||
230 |||||||
231 CAGATTAAGTGATCTCTGCAATTTCTGCAACAGAGAGGAAATAGATGAGAAATGAAGAA 490
232 |||||||
233 |||||||
234 |||||||
235 |||||||
236 |||||||
237 |||||||
238 CAAATCAAAAGTATCTCTGCAATTTCTGCAACAGAGAGGAAATAGATGAGAAATGAAGAA 496
239 |||||||
240 |||||||
241 |||||||
242 |||||||
243 |||||||
244 |||||||
245 |||||||
246 |||||||
247 |||||||
248 |||||||
249 |||||||
250 |||||||
251 ATTACCAACACTTCTTTGCAACAGTCACTGATTTGGTCCCGCAGAGAGAAAGCTATTGAA 550
252 |||||||
253 |||||||
254 |||||||
255 |||||||
256 |||||||
257 |||||||
258 ATTACTACAACTTCTTTGCAACAGTCACTGATTTGGTCCCGCAGAGAGAAAGCTATTGAA 556
259 |||||||
260 |||||||
261 |||||||
262 |||||||
263 |||||||
264 |||||||
265 |||||||
266 |||||||
267 |||||||
268 |||||||
269 |||||||
270 |||||||
271 AACCAGAGATTTTCTTAAAAATTCAAAAATCTGAGAGATTTGATACGTGACCATGTAT 610
272 |||||||
273 |||||||
274 |||||||
275 |||||||
276 |||||||
277 |||||||
278 AACCAGAGATTTTCTTAAAAATTCAAAAATCTGAGAGATTTGATACGTGACCATGTAT 616
279 |||||||
280 |||||||
281 |||||||
282 |||||||
283 |||||||
284 |||||||
285 |||||||
286 |||||||
287 |||||||
288 |||||||
289 |||||||
290 |||||||
291 TGATCAATCCCACTCTAATATACAGTTCTGAGTTACAGAGCTTTGAGGAGGAGGAGAA 670
292 |||||||
293 |||||||
294 |||||||
295 |||||||
296 |||||||
297 |||||||
298 TGGATCAATCCCACTCTAATATACAGTTCTGAGTTACAGAGCTTTGAGGAGGAGGAGAA 676
299 |||||||
300 |||||||
301 |||||||
302 |||||||
303 |||||||
304 |||||||
305 |||||||
306 |||||||
307 |||||||
308 |||||||
309 |||||||
310 |||||||
311 GA 672
312 ||
313 ||
314 ||
315 ||
316 ||
317 GA 678
318 ||
319 ||
320 ||
321 ||
322 ||
323 ||
324 ||
325 ||
326 ||
327 ||
328 ||
329 ||
330 ||
331 ||
332 ||
333 ||
334 ||
335 ||
336 ||
337 ||
338 ||
339 ||
340 ||
341 ||
342 ||
343 ||
344 ||
345 ||
346 ||
347 ||
348 ||
349 ||
350 ||
351 ||
352 ||
353 ||
354 ||
355 ||
356 ||
357 ||
358 ||
359 ||
360 ||
361 ||
362 ||
363 ||
364 ||
365 ||
366 ||
367 ||
368 ||
369 ||
370 ||
371 ||
372 ||
373 ||
374 ||
375 ||
376 ||
377 ||
378 ||
379 ||
380 ||
381 ||
382 ||
383 ||
384 ||
385 ||
386 ||
387 ||
388 ||
389 ||
390 ||
391 ||
392 ||
393 ||
394 ||
395 ||
396 ||
397 ||
398 ||
399 ||
400 ||
401 ||
402 ||
403 ||
404 ||
405 ||
406 ||
407 ||
408 ||
409 ||
410 ||
411 ||
412 ||
413 ||
414 ||
415 ||
416 ||
417 ||
418 ||
419 ||
420 ||
421 ||
422 ||
423 ||
424 ||
425 ||
426 ||
427 ||
428 ||
429 ||
430 ||
431 ||
432 ||
433 ||
434 ||
435 ||
436 ||
437 ||
438 ||
439 ||
440 ||
441 ||
442 ||
443 ||
444 ||
445 ||
446 ||
447 ||
448 ||
449 ||
450 ||
451 ||
452 ||
453 ||
454 ||
455 ||
456 ||
457 ||
458 ||
459 ||
460 ||
461 ||
462 ||
463 ||
464 ||
465 ||
466 ||
467 ||
468 ||
469 ||
470 ||
471 ||
472 ||
473 ||
474 ||
475 ||
476 ||
477 ||
478 ||
479 ||
480 ||
481 ||
482 ||
483 ||
484 ||
485 ||
486 ||
487 ||
488 ||
489 ||
490 ||
491 ||
492 ||
493 ||
494 ||
495 ||
496 ||
497 ||
498 ||
499 ||
500 ||
501 ||
502 ||
503 ||
504 ||
505 ||
506 ||
507 ||
508 ||
509 ||
510 ||
511 ||
512 ||
513 ||
514 ||
515 ||
516 ||
517 ||
518 ||
519 ||
520 ||
521 ||
522 ||
523 ||
524 ||
525 ||
526 ||
527 ||
528 ||
529 ||
530 ||
531 ||
532 ||
533 ||
534 ||
535 ||
536 ||
537 ||
538 ||
539 ||
540 ||
541 ||
542 ||
543 ||
544 ||
545 ||
546 ||
547 ||
548 ||
549 ||
550 ||
551 ||
552 ||
553 ||
554 ||
555 ||
556 ||
557 ||
558 ||
559 ||
560 ||
561 ||
562 ||
563 ||
564 ||
565 ||
566 ||
567 ||
568 ||
569 ||
570 ||
571 ||
572 ||
573 ||
574 ||
575 ||
576 ||
577 ||
578 ||
579 ||
580 ||
581 ||
582 ||
583 ||
584 ||
585 ||
586 ||
587 ||
588 ||
589 ||
590 ||
591 ||
592 ||
593 ||
594 ||
595 ||
596 ||
597 ||
598 ||
599 ||
600 ||
601 ||
602 ||
603 ||
604 ||
605 ||
606 ||
607 ||
608 ||
609 ||
610 ||
611 ||
612 ||
613 ||
614 ||
615 ||
616 ||
617 ||
618 ||
619 ||
620 ||
621 ||
622 ||
623 ||
624 ||
625 ||
626 ||
627 ||
628 ||
629 ||
630 ||
631 ||
632 ||
633 ||
634 ||
635 ||
636 ||
637 ||
638 ||
639 ||
640 ||
641 ||
642 ||
643 ||
644 ||
645 ||
646 ||
647 ||
648 ||
649 ||
650 ||
651 ||
652 ||
653 ||
654 ||
655 ||
656 ||
657 ||
658 ||
659 ||
660 ||
661 ||
662 ||
663 ||
664 ||
665 ||
666 ||
667 ||
668 ||
669 ||
670 ||
671 ||
672 ||
673 ||
674 ||
675 ||
676 ||
677 ||
678 ||
679 ||
680 ||
681 ||
682 ||
683 ||
684 ||
685 ||
686 ||
687 ||
688 ||
689 ||
690 ||
691 ||
692 ||
693 ||
694 ||
695 ||
696 ||
697 ||
698 ||
699 ||
700 ||
701 ||
702 ||
703 ||
704 ||
705 ||
706 ||
707 ||
708 ||
709 ||
710 ||
711 ||
712 ||
713 ||
714 ||
715 ||
716 ||
717 ||
718 ||
719 ||
720 ||
721 ||
722 ||
723 ||
724 ||
725 ||
726 ||
727 ||
728 ||
729 ||
730 ||
731 ||
732 ||
733 ||
734 ||
735 ||
736 ||
737 ||
738 ||
739 ||
740 ||
741 ||
742 ||
743 ||
744 ||
745 ||
746 ||
747 ||
748 ||
749 ||
750 ||
751 ||
752 ||
753 ||
754 ||
755 ||
756 ||
757 ||
758 ||
759 ||
760 ||
761 ||
762 ||
763 ||
764 ||
765 ||
766 ||
767 ||
768 ||
769 ||
770 ||
771 ||
772 ||
773 ||
774 ||
775 ||
776 ||
777 ||
778 ||
779 ||
780 ||
781 ||
782 ||
783 ||
784 ||
785 ||
786 ||
787 ||
788 ||
789 ||
790 ||
791 ||
792 ||
793 ||
794 ||
795 ||
796 ||
797 ||
798 ||
799 ||
800 ||
801 ||
802 ||
803 ||
804 ||
805 ||
806 ||
807 ||
808 ||
809 ||
810 ||
811 ||
812 ||
813 ||
814 ||
815 ||
816 ||
817 ||
818 ||
819 ||
820 ||
821 ||
822 ||
823 ||
824 ||
825 ||
826 ||
827 ||
828 ||
829 ||
830 ||
831 ||
832 ||
833 ||
834 ||
835 ||
836 ||
837 ||
838 ||
839 ||
840 ||
841 ||
842 ||
843 ||
844 ||
845 ||
846 ||
847 ||
848 ||
849 ||
850 ||
851 ||
852 ||
853 ||
854 ||
855 ||
856 ||
857 ||
858 ||
859 ||
860 ||
861 ||
862 ||
863 ||
864 ||
865 ||
866 ||
867 ||
868 ||
869 ||
870 ||
871 ||
872 ||
873 ||
874 ||
875 ||
876 ||
877 ||
878 ||
879 ||
880 ||
881 ||
882 ||
883 ||
884 ||
885 ||
886 ||
887 ||
888 ||
889 ||
890 ||
891 ||
892 ||
893 ||
894 ||
895 ||
896 ||
897 ||
898 ||
899 ||
900 ||
901 ||
902 ||
903 ||
904 ||
905 ||
906 ||
907 ||
908 ||
909 ||
910 ||
911 ||
912 ||
913 ||
914 ||
915 ||
916 ||
917 ||
918 ||
919 ||
920 ||
921 ||
922 ||
923 ||
924 ||
925 ||
926 ||
927 ||
928 ||
929 ||
930 ||
931 ||
932 ||
933 ||
934 ||
935 ||
936 ||
937 ||
938 ||
939 ||
940 ||
941 ||
942 ||
943 ||
944 ||
945 ||
946 ||
947 ||
948 ||
949 ||
950 ||
951 ||
952 ||
953 ||
954 ||
955 ||
956 ||
957 ||
958 ||
959 ||
960 ||
961 ||
962 ||
963 ||
964 ||
965 ||
966 ||
967 ||
968 ||
969 ||
970 ||
971 ||
972 ||
973 ||
974 ||
975 ||
976 ||
977 ||
978 ||
979 ||
980 ||
981 ||
982 ||
983 ||
984 ||
985 ||
986 ||
987 ||
988 ||
989 ||
990 ||
991 ||
992 ||
993 ||
994 ||
995 ||
996 ||
997 ||
998 ||
999 ||
1000 ||

```

```

source
1. 600
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="EICA006A12"
/tissue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus cartilage fetus"
/notes="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN
Query Match 44.4%; Score 531.8; DB 9; Length 600;
Best Local Similarity 93.7%; Pred. No. 6.5e-92; Indels 0; Gaps 0;
Matches 554; Conservative

QY 24 CAAAGCAAGGAAAGTAGTCTGTCTGAGAGACCATGCGCAAGAAATCTCCAGAGAAAT 83
DB 10 CGAAGCAGGAAAGAGTAGTCTGTCTGCGGAGAGATCATGCAAAAGAAATCTCCAGAGAACT 69
QY 84 GTGAAGACTGTCAATCTTAAATGAGAGAGCTTTTAAATCCGAAGAAATATGTAATCAC 143
DB 70 GTGAGGACTGTCAATCTTAAATGAGAGAGCTTTTAAATCCGAAGAGATATGTAATCAC 129
QY 144 TTAAGATTTGTGAGACTGGTGTGTTGGTATCTCTGCCCTAACTCTAAATGTCCTGTTTGGG 203
DB 130 TTAAGATTTGTGAGACTGGTGTGTTGGTATCTCTGCCCTAACTCTAAATGTCCTGTTTGGG 189
QY 204 GGAGCAAGCACTTCTGGCCGGAGAGTACCCAAAAAGCCCTATGACATGGAGCACTTTCT 263
DB 190 GGAGTAAAGCACTTCTGGCCCTGAGACACCCAAAAACATATGACATGGAGCACTTTCT 249
QY 264 ACAGCAATGGAGAGAAAGAGATTTAATGAGAAATGATCTCTGTGACGAGAACTGAA 323
DB 250 ACAGCAATGGAGAGAAAGAGATTTAATGAGAAATGATCTCTGTGACGAGAACTGAA 309
QY 324 TATTCAAGCCGAAATGGCACTGATGAAACATTTGGAAGTGACAGCTTTTAAAAACGGAT 383
DB 310 TATTCAAGCCGAAATGGCACTGATGAAACATTTGGAAGTGACAGCTTTTAAAAACGGAT 369
QY 384 ACATGCGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAAAATCAGATTAAGTGA 443
DB 370 ACATGCGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAAAATCAGATTAAGTGA 429
QY 444 TTCTGGAATTTCTGAACCAAGAGAGAAATAGATGAGAAATGAGAAATTAACCAACTT 503
DB 430 TTCTGGAATTTCTGAACCAAGAGAGAAATAGATGAGAAATGAGAAATTAACCAACTT 489
QY 504 TCCTTGAACAGTCAGTATTTGGTCCCGCAGAGAAAGCCCTATTGAAAAACCGAGATTTTC 563
DB 490 TCCTTGAACAGTCAGTATTTGGTCCCGCAGAGAGAAACCTATTGAAAAACCGAGATTTTC 549
QY 564 TTAATAATTCCTCAAAATTTCTGAGATTTGTGATACCGTACCATGTATTGGA 614
DB 550 TTAATAATTCCTCAAAATTTCTGAGATTTGTGATACCGTACCATGTATTGGA 600

RESULT 24
CD773806 639 bp mRNA linear EST 02-JUL-2003
LOCUS AGENCOURT 14719761 NIH MGC 190 Mus musculus cDNA clone
DEFINITION IMAGE:30504556 5', mRNA sequence.
CD773806
VERSION CD773806.1 GI:32432308
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 639)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabbs-r@mail.nih.gov  
Tissue procurement: Yoshihiko Yamada, Takashi Nakamura, NIDCR  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCM232 row: d column: 05  
High quality sequence stop: 529.  
Location/Qualifiers  
1. .639  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 190"  
/note="Organ: Pooled - Molar: Vector: pDNR-LIB; Site 1:  
SfiI (ggccattatggcc); Site 2: SfiI (ggccctctggcc);  
Non-normalized full-length enriched library 5' and 3'  
adaptors were used in cloning as follows: 5' adaptor  
5'-ATTCTAGAGCGCGCGCGGCACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.71  
kb (range 0.5-3.0 kb). 15/15 colonies contained inserts by  
PCR. This library was enriched for full-length clones and  
was constructed by Clontech Laboratories (Palo Alto, CA)  
Corp."

FEATURES  
source

RESULT 25  
BI694698  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI694698 603347638F1 NCI\_CGAP\_Mam2 Mus musculus cDNA clone linear EST 18-SEP-2001  
mRNA sequence.

BI694698  
EST. BI694698.1 GI:15657327

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1955 row: k column: 10  
High quality sequence start: 8  
High quality sequence stop: 732.  
Location/Qualifiers

1. .735

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N-3"

/db\_xref="taxon:10090"

/clone="IMAGE:5375361"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="5 months"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Mam2"

/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

FEATURES  
source

Query Match 44.3%; Score 530.4; DB 14; Length 639;  
Best Local Similarity 91.5%; Pred. No. 1.2e-91;  
Matches 572; Conservative 0; Mismatches 52; Indels 1; Gaps 1;  
20 CTCTCAAGCAAGCAAGAAAGTACTGTGTGCTGAGAGACCATTGGCAAGAAATCCTCCAGAG 79  
7 CTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 66  
80 AATTGTGAAGACTGTCACTTCTAAATGCAAGACTTTAAATCCAAAGAAATATGTAAA 139  
67 AACTGTGAGGGCTGTCACTTCTAAATGCAAGACTCTGAAATCTAAGAGATATGTAAA 126  
140 TCACTTAAGATTGTGTGACTGGTGTGTGATCTGTGCGCCCTTAATCTTAATGTCTCTTT 199  
127 TCACTGAAGATTGTGGACTAGTGTGTGATCTGTGCGCCCTTAATCTTAATGTCTCTTT 186  
200 TGGGGGAGCAAGCACTTCTGCGCGGAGGTACCCAAAGAAAGCCTATGATGAGGACACT 259  
187 TGGGGGAGCAAGCACTTCTGCGCGGAGGTATCCAAAGAAACCTATGATGAGGACACT 246  
260 TTCTACAGCAATGGAGAGAGAAAGATTACATGGAAATTTGATCTGTGACCAAGAACT 319  
247 TTCTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 306  
320 GAAATATTGAGAGCGGAATGCGCTGATGAACCAATGGAAGTGCACGCTTTAAAAAC 379  
307 GAAATATTGAGAGTGGAAATGCGCTGATGAACCAATGGAAGTGCACGCTTTAAAAAT 366  
380 GGATACACTGGCACTACTTCTGTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAA 439  
367 GGATACACTGGCACTACTTGTAGTCTTCAAAAATGCTTTTAAAACTCAATCAAA 426  
440 GTGATTCCTGAATTTCTGAACCAAGAGGAATATAGATGAGAAATGAAGAAATTCACCA 499  
427 GTGATTCCTGAATTTCTGAACCAAGAGGAATATAGATGAGAAATGAAGAAATTCACCA 486  
500 ACTTCTTTTGAACAGTCAGTGATTGGTGGTCCCGAGCAAGAAAGCCTATTGAAACCCGAGAT 559

## ORIGIN

Db 487 ACTTCTTTTGAACAGTCAGTGATTGGTTCGCGCAAGAGCCTATTGAAAAACAGAGA - 545  
QY 560 TTTCTTAAAAATTCCAAATTCCTGGAGATTGTGATAAGCTGACCATGTATTGGATCAAT 619  
Db 546 CTTCTCGAAAAATCTAAATTCCTGGAGATTTCGGATAATGTGACCATGTACTGGATCAAT 605  
QY 620 CCACCTCTAATATCAGTTTCTGAGT 644  
Db 606 CCCACTCTAATAGCAGTTTCAGAA 630

## ORIGIN

Query Match 41.8%; Score 500.6; DB 12; Length 735;  
Best Local Similarity 88.7%; Pred. No. 5.7e-86;  
Matches 611; Conservative 0; Mismatches 69; Indels 9; Gaps 6;  
QY 246 ACATGGAGCACACTTCTTCTACA--GCATGGAGAGAGAGAAAGATTTCATGGAAATTG 302  
Db 47 ACATGGAGCACACTTCTTCTACATGCAACGCGCGAGAGAAAGAGATTTCATGGAAATTG 106  
QY 303 ATCTCTGTACCAAGAACTGAAATATTACAGAGCGGAATGCGCACTGATGAAACATTGGAG 362  
Db 107 ATCCCATACCAAGACAGAAATATTTCAGAAATGGCAATGGCACTGATGAAACATTGGA 166  
QY 363 TCACGACACTTTAAAAACGGATACACTGGCATCTACTTCTGTGGTCTTCAAAAATGTTTAA 422  
Db 167 TCCATGACTTTAAAAATGAATACATGGCATCTACTTTGTAGGTCTTCAAAAATGCTTTA 226

QY 423 TCAGAACTCAGATTAAGTCTGATTTCTGACACAGAGAGAAATAGATGAGA 482  
 Db 227 TTAAGACTCAATCAAGTGAATCTGAAATTTCTGAACACAGAGAGAAATAGATGAGA 286  
 QY 483 ATGAAGAAATACCAACACTTTCTTTGAACAGTCTGATTTGGTCCGACAGAGAAAGC 542  
 Db 287 ATGAAGAAATACCAACACTTTCTTTGAACAGTCTGATTTGGTCCGACAGAGAAAGC 346  
 QY 543 CTATTGAAACCGAGATTTCTTTAAATTTCAAAATTTCTGAGATTTGTGATACGTGA 602  
 Db 347 CTATTGAAACCGAGATTTCTTTAAATTTCTGAGATTTGTGATACGTGA 406  
 QY 603 CCATGATTTGATCAATCCACTCTAAATACATTTCTGAGTTACAGACTTTGAGAGAG 662  
 Db 407 CCATGATTTGATCAATCCACTCTAAATACAGTTTTCAGAAATACAGACTTTGAGAGAG 466  
 QY 663 AGGAGAGAGATCTTCACTTTCTGACCAAGAAAGAGGATTTGAACAA--AATGAACA 720  
 Db 467 ACGGTGAAGATCTTCACTTTCTGACAGCTGAAAGAGAGGATTTGACCAAGCAATGAGCA 526  
 QY 721 GTGGGTGTCCTCAAGTGAAGTAGA--GAAGACCCGTCAGCCAGACAGCAAGCAAGTGAG 778  
 Db 527 ATGGGTGTCCTCAAGTGAAGTAGA--GAAGACCCGTCAGCCAGACAGCAAGCAAGTGAG 586  
 QY 779 GAAGACCTTCCATTAATCACTATCTGAAAT--GGAATAGAAATTTGATCCCATGCTGGA 837  
 Db 587 GAAGACCTTCCATTAATCACTATCTGAAAT--GGAATAGAAATTTGATCCCATGCTGGA 646  
 QY 838 TGAGAGAGGTATTGTTGTTTACTGCGCTGAGGCAACCGCTATGCGCGCGCTCTG 897  
 Db 647 TGAGAGAGGTATTGTTGTTTACTGCGCTGAGGCAACCGCTATGCGCGCGCTCTG 705  
 QY 898 TGAACCTTTACTAGGCTACTACCCATATC 926  
 Db 706 TGAACCTTTACTAGGCTACTACCCATACC 734

RESULT 26  
 BI535437 557 bp mRNA linear EST 30-AUG-2001  
 LOCUS 398969 MARC 4BOV Bos taurus cdNA 5', mRNA sequence.  
 DEFINITION BI535437  
 ACCESSION BI535437  
 VERSION BI535437.1 GI:15376545  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bos.  
 1 (bases 1 to 557)  
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,  
 Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A.,  
 Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,  
 Quackenbush,J. and Keele,J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cdNA  
 libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 21180013  
 11282978  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980504.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR PRIMERs  
 FORWARD: AGGAAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 127 row: C column: 3

FEATURES  
 source  
 Seq primer: ATTTAGGTGACACTATAG.  
 Location/Qualifiers  
 1..557  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="MARC 4BOV"  
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
 Library made from pooled tissue from day 20 and day 40  
 embryos."

ORIGIN  
 Query Match 41.7%; Score 498.6; DB 12; Length 557;  
 Best Local Similarity 93.9%; Pred. No. 1.5e-85;  
 Matches 519; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
 QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGCAAGAGAGTACTGTGTCTGAGAGACCAT 60  
 Db 5 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGCAAGAGAGTACTGTGTCTGAGAGATCAT 64  
 QY 61 GGCAAGAAATCCCTCCAGAGAAATGTGAAGACTGTCACTTCTAAATGCAAGACTTTTAA 120  
 Db 65 GGCAAGAAATCCCTCCAGAGAACTGTGAGGACTGTCACTTCTAAATGCAAGACTTTTAA 124  
 QY 121 ATCCAGAAATATATGAAATCACTTAAAGATTTGTGACTGCTGTTGGTATCTCGGCCCT 180  
 Db 125 ATCCAGAAATATGAAATCACTTAAAGATTTGTGAAATTTGGTATCTCTCGGCCCT 184  
 QY 181 AACTCTAATTTGCTGTTTGGGGGAGCAAGCACTTCTGCGGAGGTACCCAAAAAAGC 240  
 Db 185 AACTCTAATTTGCTGTTTGGGGGAGTAAAGCACTTCTGCGCTGAGACACCCAAAAAAGC 244  
 QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAT 300  
 Db 245 ATATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAT 304  
 QY 301 TGATCTGTGACAGAACTGAAATATTCAGAACGGGAAATGGCACTGATGAACATTTGA 360  
 Db 305 TGATCTGTGACAGAACTGAAATATTCAGAACGGGAAATGGCACTGATGAACATTTGA 364  
 QY 361 AGTGACGACTTTAAAGAACGATACACTGCTGCTGCTTCTCAAAAATGTTT 420  
 Db 365 AGTACATGACTTTAAAGAACGATACACTGCTGCTTCTTGTAGTCTTCAAAAATGTTT 424  
 QY 421 TATCAAACTCAGATTAAGTGAATTTCTGAATTTCTGAACAGAGAGAAATAGATGA 480  
 Db 425 TATCAAACTCAGATTAAGTGAATTTCTGAATTTCTGAACAGAGAGAAATAGATGA 484  
 QY 481 GAATGAAGAAATACCAACTTTCTTTGAACAGTCACTGATTTGGTCCGACAGAGAAA 540  
 Db 485 GAATGAAGAAATACCAACTTTCTTTGAACAGTCACTGATTTGGTCCGACAGAGAAA 544  
 QY 541 GCCTATTGAAAC 553  
 Db 545 ACCTATTGAAAC 557

RESULT 27  
 AW743952 588 bp mRNA linear EST 27-APR-2000  
 LOCUS ur24h02.y1 Soares mouse NMHP Mus musculus cdNA clone IMAGE:3025299  
 DEFINITION 5', similar to TR:Q9YI63 Q9YI63 CHONDROMODULIN-I PRECURSOR. ; mRNA  
 sequence.  
 ACCESSION AW743952  
 VERSION AW743952.1 GI:7655775  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 588)



```

Db 479 ACCGAGACAGAAAATTTCAGAAAGTGGAATGGCACTGATGAAACATTGGAAGTCCATGAC 538
QY 371 TTTAAAAACGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTATCAAAACT 430
Db 539 TTTAAANTGATACACTGGCATCTACTTTGTAGGTCTTCAAAAATGCTTTTATTAATACT 598
QY 431 CAGATTAAGTGATTCCTGAATTTCTGAACACAGAGAGGAATAGATGAGAAATGAAGAA 490
Db 599 CAAATCAAAGTGATTCCTGAATTTCTGAACACAGAGGAAGAAATAGATGAGAAATGAAGAA 658
QY 491 ATTACCAACACTTCTTTTG-AACAGTCAGTGATTTGGGTCCCGACGAG-AAAGCCCTATTG 548
Db 659 ATTACTAACACTTCTTTTGAAACAGTCAGTGATTTGGGTTCCTCCGACAGAAAGCCCTATTG 718
QY 549 AAAACCGAGATTTTCT-TAAAAATTCAAAATTCGGAGATTTGTGATAACGTGACCATG 607
Db 719 AAAACAGAGACTTCTCGAAAAAATTTCTAAAATTTCTGGAGATTTGGGATAATGTGACCATG 778
QY 608 TATTGATCAATCCCACTCT 627
Db 779 TACTGGATCAATCCCACTCT 798

RESULT 29
CB440637 610 bp mRNA linear EST 25-MAR-2003
LOCUS 690932 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION
ACCESSION CB440637
VERSION CB440637.1 GI:29226859
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 610)
Smith, J.F.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keefe, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329
Plate: F018047 row: F column: 8
Seq primer: GTAATACGACTCACTATAGG.
FEATURES
Location/Qualifiers
1..610
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
ORIGIN
Query Match 40.6%; Score 485.8; DB 14; Length 610;
Best Local Similarity 90.2%; Pred. No. 4.1e-83;
Matches 544; Conservative 0; Mismatches 52; Indels 7; Gaps 2;
QY 535 AGAAAAGCTATTGAAACCGAGATTTCTTAAAAATTCCAAATTCGGAGATTTGGA 594
Db 8 AGAAAACCTATTGAAACCGAGACTTTCTTAAAAATTCCAAATTCGGAGATTTGGA 67

```

```

QY 595 TAACGTGACCATGATTTGGATCAATCCCACTCTAATATATCACTTTCTGAGTTACAGACTT 654
Db 68 TAATGTGACCATGATTTGGATCAATCCCACTCTAATAGCACTTTTCAGAGTTACAGACTT 127
QY 655 TGAGGAGGAGGAGAGATCTTCACTTTCTCTGCCAAACGAAAAAAGGATTTGAACAAA 714
Db 128 TGGAGGATGAGTGAAGACCTTCACTTTCTCTACAGTGAAGAAAAAGGCAATTTGAACAAA 187
QY 715 TGAACAGTGGGTGGTCCCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771
Db 188 CGAGCAGTGGGTGGTCCCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247
QY 772 AAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 831
Db 248 AAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 307
QY 832 GCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891
Db 308 GTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367
QY 892 CGTCTGTGAACCTTTTACTAGGCTACTACCCATATCCATCTCTCTACCAAGGAGAGAGAG 951
Db 368 CGTCTGTGAACCTTTTACTAGGCTACTACCCATATCCATCTCTCTACCAAGGAGAGAGAG 427
QY 952 CATCTGTCTGTCTATCATGCTTGTAACTGTGGGTGGCCCGCATGCTCTCTACCAAGGAGAGAG 1011
Db 428 TATCTGTCTGTCTATCATGCTTGTAACTGTGGGTGGCCCGCATGCTCTCTACCAAGGAGAGAG 487
QY 1012 ATAGGAGGTTCGAGCTCAAACTGCTTAACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1067
Db 488 ATAGCAAGTCTGAGCTCGAGTCTTAACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 547
QY 1068 TATTCAATGAATTTCTGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1127
Db 548 TATTCAATGAATTTCTGCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 607
QY 1128 CTT 1130
Db 608 CTT 610

RESULT 30
BI304069/c 636 bp mRNA linear EST 20-JUL-2001
LOCUS UI-R-DR0-cjc-o-14-0-UI.s1 UI-R-DR0 Rattus norvegicus cDNA clone
DEFINITION UI-R-DR0-cjc-o-14-0-UI 3', mRNA sequence.
ACCESSION BI304069
VERSION BI304069.1 GI:14980349
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 636)
Bonald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE 889548
PUBMED
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site

```

atch	39.7%;	Score 474.8;	DB 9;	Length 479;
cal Similarity	99.6%;	Pred. 0. 5.7e-81;		
476; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
699	AAGGGATTGAA	CAAAAATGAA	CAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTC	758
479	AAGGGATTGAA	CAAAAATGAA	CAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTC	420
759	ACGCCAGACAG	CAAGTGAAGAACTT	CCAAATAAATGACTATCTCGAAATGGAATAG	818
419	ACGCCAGACAG	CAAGTGAAGAACTT	CCAAATAAATGACTATCTCGAAATGGAATAG	360
819	AATTTGATCCCATG	CTGGATCAGAGAGGTTAT	TGTTGTTATTTACTGCCGTGAGGCAACC	878
359	AATTTGATCCCATG	CTGGATCAGAGAGGTTAT	TGTTGTTATTTACTGCCGTGAGGCAACC	300
879	GCTATTGGCGCGGT	CTGTGAACTTTACTAGGCTAC	CCCATACCATCTGCTACC	938
299	GCTATTGGCGCGGT	CTGTGAACTTTACTAGGCTAC	CCCATACCATCTGCTACC	240

Qy	939	AAGGAGAGCAGTCACTGCTGTCATCATGCTTGTAACTGCTGGTGGCCGCGATGC	998	Db	475	AACGAGCAGTGGGTGGTCCCTCAAGTGAAGTGAGAGACCCCTCACACAGCAAGCA	416
Db	239	AAGGAGGACGAGTCACTGCTGCTCATCATGCTTGTAACTGCTGGTGGCCGCGATGC	180	Qy	770	GCAAGTGAAGAGAACTTCCAAATAATGACTATATGAAATGGAATGAAATTTGATCCC	829
Qy	999	TGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATAATA	1058	Db	415	GCAAGTGAAGAGAACTTCCCAATCAATGACTATATGAAATGGAATGAAATTCGACTCC	356
Db	179	TGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATAATA	120	Qy	830	ATGCTGGATGAGAGAGGTTATTTGTTATTTACTGCGGTCGAGGCAACCGCTATTCGCG	889
Qy	1059	AATGATGCTATTCAATGAATTTCTGCTATGAGCAATCTGCGCCCTGGTAGCGAGCTCT	1118	Db	355	ATGTTGGATGAGAGAGGTTACTGTTGTTATTTACTGCGGTCGAGGCAACCGCTATTCGCG	296
Db	119	AATGATGCTATTCAATGAATTTCTGCTATGAGCAATCTGCGCCCTGGTAGCGAGCTCT	60	Qy	890	CGGCTGCTGAGACCTTTACTAGGCTACTACCATATCCATATCTGCTCAACAGGAGGAGGA	949
Qy	1119	CCAGAACTACTGAGGTAATCTCTCTTCTCATGTTCTTAATAAACTTCTACATATCA	1176	Db	295	CGGCTGCTGAGACCTTTACTAGGTTACTACCCGATATCCATATTTGCTACCAAGGAGGCGG	236
Db	59	CCAGAACTACTGAGGTAATCTCTCTTCTCATGTTCTTAATAAACTTCTACATATCA	2	Qy	950	GTCACTCTGCTGTCATCATGCTTGTAACTGCTGGTGGTGGCCGCGCATCTGGGGAGGGTC	1009
RESULT 32				Db	235	GTTATCTGCTGTCATCATGCTTGTAACTGCTGGTGGTGGCCGCGCATCTGAGGAGGGTT	176
CB440284/c				Qy	1010	TAATAGGAGGTTGAGCTCAATGCTTAACTGCTGGTGGTGGCCGCGCATCTGAGGAGGGT	1065
LOCUS	690548	MARC 6BOV Bos taurus cDNA 3', mRNA	linear	EST 25-MAR-2003	175	TAATAGCAAGTCTGAGCTCGAGTGTCTTAACTTCTGGCATCCAAATATAAATGCAAT	116
DEFINITION	CB440284			Qy	1066	GCTATTCAATGAATTTCTGCTGCTATGAGGCAATCTGGCCCTGCTAGCCAGCTCTCCAGAA	1125
ACCESSION	CB440284			Db	115	GCTATTCAATGAATTTCTGCTGCTATGAGGCAATTTGGCTCTGCTAGCCATATCTCCGAA	56
VERSION	CB440284.1	GI:29226148		Qy	1126	TACTTGTAGGTAATTTCTCTTCTTCTTCAATGTTCTTAATAAACTTCTACATATACCAA	1180
KEYWORDS	EST.			Db	55	TGCTTGTAGGTAATTTCTCTTCTTCTTCAATGTTCTTAATAAACTTCTACACCCCTCAAAA	1
SOURCE	Bos taurus (cow)			RESULT 33			
ORGANISM	Bos taurus			LOCUS	BX104255		
REFERENCE	Smith, R.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,			DEFINITION	IMAGp9980034297	468 bp	mRNA linear EST 06-FEB-2003
AUTHORS	Wray, J.E. and Keele, W.W.			ACCESSION	BX104255		Soares fetal heart NbhH19W Homo sapiens cDNA clone
TITLE	A second set of bovine ESTs from pooled-tissue normalized libraries			VERSION	BX104255		IMAGp9980034297; IMAGE:1692506, mRNA sequence.
JOURNAL	Unpublished (2003)			KEYWORDS	EST.		
COMMENT	Contact: Smith TPL			SOURCE	Homo sapiens (human)		
	USDA, ARS, US Meat Animal Research Center			ORGANISM	Homo sapiens		
	PO Box 166, Clay Center, NE 68933-0166, USA				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Tel: 402 762 4366				Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	Fax: 402 762 4390				1 (bases 1 to 468)		
	Email: smith@mail.marc.usda.gov				Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,		
	Single pass sequencing. Bases called with phred v0.020425.c and				Radelof, U., Schneider, D. and Korn, B.		
	trimmed with the aid of the trim_alt option. Vector identified with				Human Unigeneset - RZPD3		
	cross_match v0.990329				Unpublished (2003)		
	Plate: FQY6047 row: F column: 8				Contact: Ina Rofls		
	Seq primer: TAGAGCGCAGTCGAGG.				RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH		
FEATURES	Location/Qualifiers				Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany		
source	1..595				RZPD; IMAGp9980034297.		
	/organism="Bos taurus"				RZPD LIB: I.M.A.G.E. cDNA Clone Collection;		
	/mol_type="mRNA"				Human Unigeneset - RZPD3 (RZPD LIB No.972)		
	/db_xref="taxon:9913"				http://www.rzpd.de/CloneCards/cgi-		
	/tissue_type="pooled"				bin/showlib.pl.cgi/responderlibNo=972 Contact: Ina Rofls		
	/lab_host="DH10B"				RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH		
	/clone_lib="MARC 6BOV"				Heubnerweg 6, D-14059 Berlin, Germany		
	/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;				Tel: +49 30 32639 101		
	Library made with RNA pooled from multiple tissues				Fax: +49 30 32639 111		
	including liver, lung, hypothalamus, pituitary, and				www.rzpd.de		
	placenta/endometrium."				This clone is available royalty-free from RZPD;		
ORIGIN					contact RZPD (clone@rzpd.de) for further information. Seq primer:		
Query Match	39.5%; Score 473; DB 14; Length 595;				M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.		
Best Local Similarity	89.6%; Pred. No. 1.2e-80;				Location/Qualifiers		
Matches 533; Conservative	0; Mismatches 55; Indels 7; Gaps 2;				1..468		
Qy	593	GATACGTCGACCATGATTTGATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAGAC	652	source	/organism="Homo sapiens"		
Db	595	GATAATGTGACCATGATTTGATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAGAC	536		/mol_type="mRNA"		
Qy	653	TTTGAGGAGGAGGAGAGATTTTCACTTCTCCCAACGAAAAAAGGAGGATTTGAACAA	712		/db_xref="taxon:9606"		
Db	535	TTTGAGGAGGATGTTGGAAGACCTTCACCTTCTCCAGTGAAAAAAGGATTTGAACAA	476		/clone="IMAGp9980034297 ; IMAGE:1692506"		
Qy	713	AATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCAGCCAGACA	769		/sex="unknown"		
					/dev_stage="19 weeks"		



```

/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBh19W"
/note="Organ: heart; Vector: pTV732 (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo (dC)T primer [5'
TGTTTACCAATCTCAAGTGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pTV73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fátima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBh19W."

```

## ORIGIN

Query Match	39.1%;	Score 468;	DB 13;	Length 468;
Best Local Similarity	100.0%;	Pred. No. 1.2e-79;		
Matches 468;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	633	CAGTTTCTGAGTTACAAGACTTTTGAGGAGGAGGAGAGATCTTCACTTTCTCTCCCAACG	692	
Db				
1		CAGTTTCTGAGTTACAAGACTTTTGAGGAGGAGGAGAGATCTTCACTTTCTCTCCCAACG	60	
QY	693	AAAAAAAAGGATTTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAAGA	752	
Db				
61		AAAAAAAAGGATTTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAAGA	120	
QY	753	CCGTCACGCCACAGCAAGTGAAGAAACTTCCAATAAATGACTATATCTGAAAATG	812	
Db				
121		CCGTCACGCCACAGCAAGTGAAGAAACTTCCAATAAATGACTATATCTGAAAATG	180	
QY	813	GAATAGAAATTGATCCATGCTCGATGAGAGAGGTTATTGTTGTTATTACTGCGTCGAG	872	
Db				
181		GAATAGAAATTGATCCATGCTCGATGAGAGAGGTTATTGTTGTTATTACTGCGTCGAG	240	
QY	873	GC AACCGCTATTGCCGCGCGTCTGTGAACCTTTACTAGCTACTACCCATATCCATACT	932	
Db				
241		GC AACCGCTATTGCCGCGCGTCTGTGAACCTTTACTAGCTACTACCCATATCCATACT	300	
QY	933	GCTACCAAGAGGACGAGTCACTGTCTGTCTATCATGCTCTCTTAAGTGGTGGTGGCCC	992	
Db				
301		GCTACCAAGAGGACGAGTCACTGTCTGTCTATCATGCTCTCTTAAGTGGTGGTGGCCC	360	
QY	993	GCATGCTGGGAGGGGTCTAATAGGAGGTTTGAGSCTCAAAATGCTTAAACTGCTGGCAACAT	1052	
Db				
361		GCATGCTGGGAGGGTCTAATAGGAGGTTTGAGSCTCAAAATGCTTAAACTGCTGGCAACAT	420	
QY	1053	ATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCAATCGG	1100	
Db				
421		ATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCAATCGG	468	

RESULT 34

REFSeq: 34  
 BF121576  
 LOCUS  
 DEFINITION  
 601756377F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3986395 5',  
 mRNA sequence.  
 BF121576  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 BF121576.1 GI:10960616  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 888)  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: LLAM9191 row: i column: 20  
High quality sequence stop: 677.  
Location/Qualifiers  
1. .888  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3986395"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam5"  
/note="Organ: Mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NTH"

## FEATURES

source

## ORIGIN

Query Match	38.7%;	Score 463.6;	DB 10;	Length 888;
Best Local Similarity	83.6%;	Pred. No. 6.5e-79;		
Matches 632;	Conservative 0;	Mismatches 114;	Indels 10;	Gaps 9;
QY	53	GAGACCATGGCAAGAATCTCCAGAGAAATGTGAAGACTGTACATTTCAAATGCGAA	112	
Db	1	GAGACCATGGCAAGAATCTCCAGAGAACTGTGAGGCTGTACATTTCAAATGCGAA	60	
QY	113	GCCTTTAAATCCAGAAAAATATGTAAATCACTTAAGATTTGTGACATGGTGTGTTGGTATC	172	
Db	61	GCCTGAAATCTAAGAGATATGTAAATCACTGAAGA-TTGTGGACTAGTG-TTGGTATC	118	
QY	173	CTGGCCCTAACTCTAAATTTGTCTGTGTTTGGGGAGCAAGCACTTCTGCGCGAGGTACCC	232	
Db	119	CTGGCCCTAACTCTAAAT- GTCCCTGTTTGGGGAGCAACACTTCTGCGCCGAGTATCC	176	
QY	233	AAAAAGCCCTATGACATGGAGCACATTTCTACGCAATGGAGAGAAGAAAGATTTC	292	
Db	177	AAGAAAACCTTATGACATGGAGCACATTTCTACAGCAACGCGAGAGAAGAAAGATTTC	236	
QY	293	ATGGAATTCATCTCTGTCACAGAACTGGAATATTTCAGAAAGCGGAATGCACCTGATGAA	352	
Db	237	ATGMAA-TGATCCCATACAGAACGAATATTTCAGAAGTGGAAATGGCACTGATGAA	295	
QY	353	ACATTTGGAAGTCACGCACTTTTAAAAACGGATACACTGGCACTCTACTTCGTGGGTCTTCAA	412	
Db	296	ACATTTGGAAGTCATGACTTTTAAAAATGGATACACTGGCATCTACTT-TGAGGTCTTCAA	354	
QY	413	AAATGTTTTATCAAACTCAGATTAAGTGAATTCCTGAATTTTCTGAACACAGAAGGAA	472	
Db	355	AAATGCTTTATTAAAACTCAAACTCAAAGTGAATTCCTGAATTTTCTGAACACAGAAGGAA	414	
QY	473	ATAGATGAGAATCAAGAAATATTACCAACTTTCTTTTGAACAGTCAGTGAATTTGGGTCCCA	532	
Db	415	ATAGATGAGATGAGAAATTTCTCAACTTTCTTTTGAACAGTCAGTGA-TTGGGTCTCC	473	
QY	533	GCAGAAAAGCCTATTGAAAACCGAGATTTTCTTTAAAAATTCGAGATTTGT	592	
Db	474	GCAGAAAAGCCTATGAAAACAGAGACTTCTTGAAAAATTTCAAATTTCTGGAGATTG-	532	
QY	593	GATAAGTGAACCATGATTGGATCAATCCCACTCTAATATCAGTTTCTCAGTTTACAAGAC	652	
Db	533	GATAATGTGACCATGATCTGAGTCAATCCCACTCTAAATAGCAGTTTTCAGAAATTTACAGAC	592	
QY	653	TTTGAGGAGGAGGAG-AGATCTTTCAC-TTTTCTGCGCAACGAAAAAAAAGGATTGAAC	710	
Db	593	TTTGAGGAGGAGCGGTGAAAGATCTTCACTTTTCTTCAACAGTGAACACAGGGGATGACC	652	

```

QY 711 AATGATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAACACCGCTCACCCAGCAAG 770
Db 653 AAGATGAGCAGAGGTGGTCCGACAGTTGAGGGGGGGAACCGCCAAACAGAAAGCAG 712
QY 771 CAAGTGAAGGAAGAACTTCCAAATAAATGACTATACGT 806
Db 713 CGAGGGAAAACCTTCATTAAAGACCAATATGGCACTG 748

RESULT 35
BF439260/c
LOCUS
DEFINITION
nab61h07.x1 Soares NSF F8.9W OT PA P S1 Homo sapiens cDNA clone
IMAGE:3270397 3' similar to TR:Q9UJG0 Q9UJG0 D44797.1 ; mRNA
sequence.
ACCESSION
BF439260
VERSION
BF439260.1 GI:11451777
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 462)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..462
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3270397"
/lab_host="PH10B"
/clone_lib="Soares NSF F8.9W OT PA P S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NB2HP pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HP-9W pool 1:
758280-760583, 772104-774407 Soares NB2HP pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 38.4%; Score 460; DB 10; Length 462;
Best Local Similarity 100.0%; Pred. No. 4e-78;
Matches 460; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 AACAGTGGGTGTCCTCAAGTGAAGTAGAGAACCCGTCACGCCAGACAAAGCAAGT 776
Db 462 AACAGTGGGTGTCCTCAAGTGAAGTAGAGAGAACCCGTCACGCCAGACAAAGCAAGT 403
QY 777 AGGAAGAACTTCCAATAAATGACTATACGTGAATGAATGAAATAGATTGATCCCATGCTGG 836
Db 402 AGGAAGAACTTCCAATAAATGACTATACGTGAATGAATGAAATAGATTGATCCCATGCTGG 343
QY 837 ATGAGAGAGTTATGTTGTTATTTACTCCGTCGAGGCAACCGCTATTGGCGCGCGTCT 896
Db 342 ATGAGAGAGTTATGTTGTTATTTACTCCGTCGAGGCAACCGCTATTGGCGCGCGTCT 283

```

```

QY 897 GTGAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGGAGGTCATCT 956
Db 282 GTGAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGGAGGTCATCT 223
QY 957 GTCGTGTCATCATCGCTTGTAACTGCTGGTGGCCCGCATGCTGGGAGGGGTCTAATAGG 1016
Db 222 GTCGTGTCATCATCGCTTGTAACTGCTGGTGGCCCGCATGCTGGGAGGGGTCTAATAGG 163
QY 1017 AGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAATAATGCTATGCTTCAATG 1076
Db 162 AGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAATAATGCTATGCTTCAATG 103
QY 1077 AATTTCTCCCTATGAGGCATCTGGCCCTGCTAGCCAGCTCTCCAGAAATTACTTCTAGGT 1136
Db 102 AATTTCTCCCTATGAGGCATCTGGCCCTGCTAGCCAGCTCTCCAGAAATTACTTCTAGGT 43
QY 1137 AATTCCTCTCTTCAATGTTCTAATAAATCTTACATTATCA 1176
Db 42 AATTCCTCTCTTCAATGTTCTAATAAATCTTACATTATCA 3

T12179 A533F Heart Homo sapiens cDNA clone A533 similar to Chondromodulin,
mRNA sequence.
T12179
T12179.1 GI:596883
ACCESSION
T12179
VERSION
T12179.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 465)
Liew, C.C., Hwang, D.M., Fung, Y.W., Laurensen, C., Cukerman, E.,
Tsui, S. and Lee, C.Y.
A catalogue of genes in the cardiovascular system as identified by
expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 91, 10645-10649 (1994)
95024171
MEDLINE
7938007
PUBMED
COMMENT
Other ESTs: A533R
Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6177750995
Email: cliw@rics.bwh.harvard.edu
Similar to bovine chondromodulin
Seq primer: GGTGGCAGCACTCTCTGAGCC.
FEATURES
Location/Qualifiers
1..465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="A533"
/lab_host="E.coli Y1090"
/clone_lib="Heart"
/note="Vector: Lambda gt11; Site 1: EcoRI; Site 2: EcoRI"
ORIGIN
Query Match 36.4%; Score 435.8; DB 14; Length 465;
Best Local Similarity 99.1%; Pred. No. 1.7e-73;
Matches 459; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 451 ATTTCTGACCAAGAGGAAATAGATGAATGAAGAAATACCAACAATTTCTTTTGA 510
Db 461 AATTTCTGAAACCAAGAGGAAATAGATGAATGAAGAAATACCAACAATTTCTTTTGA 403
QY 511 ACAGTCAGTGTATTTGGTCCAGCAAAAAGCCCTATTGAAACCGAGATTTCTTAAAAA 570

```

```
Db 402 ACAGTCAGTGATTTGGGTCCCGACGAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAA 343
QY 571 TTCAAAAATTCGGAGATTGTGATAAAGTGCACCATGATTGATCAATCCACACTTAAT 630
Db 342 TTCAAAAATTCGGAGATTGTGATAAAGTGCACCATGATTGATCAATCCACACTTAAT 283
QY 631 ATCAGTTTCTGAGTTACAAGACTTTGAGAGAGGAGGAGAAAGATCTTCACTTTCTTCGCCAA 690
Db 282 ATCAGTCTCTGAGTTACAAGACTTTGAGAGAGGAGGAGAAAGATCTTCACTTTCTTCGCCAA 223
QY 691 CGAAAAAAGGGATTGAACAAAATGACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAA 750
Db 222 CGAAAAAAGGGATTGAACAAAATGACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAA 163
QY 751 GACCGTCAACCGACGACGAGCAAGTGAAGAACTTCCAATAAATGACTACTAGAAA 810
Db 162 GACCGTCAACCGACGACGAGCAAGTGAAGAACTTCCAATAAATGACTACTAGAAA 103
QY 811 TGGATAGAAATTTGATCCATGCTGGATGAGAGAGGTTATTGTTGATTTACTGCCGTCG 870
Db 102 TGGATAGAAATTTGAT-CCATGCTGGATGAGAGAGGTTATTGTTGATTTACTGCCGTCG 44
QY 871 AGGCAACCGCTATTGCCCGCGCTGTGACCTTTACTAGC 913
Db 43 AGGCAACCGCTATTGCCCGCGCTGTGACCTTTACTAGC 1

RESULT 37
BF123957 856 bp mRNA linear EST 24-OCT-2000
LOCUS 601759641F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4022674 5',
DEFINITION mRNA sequence.
ACCESSION BF123957
VERSION BF123957.1 GI:10962997
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9280 row: a column: 11
High quality sequence stop: 703.
FEATURES
Location/Qualifiers
1..856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4022674"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
```

ORIGIN

Query Match

35.5%; Score 424.6; DB 10; Length 856;

Best Local Similarity 89.5%; Pred. No. 1.9e-71;  
Matches 513; Conservative 0; Mismatches 54; Indels 6; Gaps 5;

```
QY 11 CTCTCAGTCTCTCAAAGCAAGGAAGTAGTACTGTGCTGAGAGACCATGGCAAGAAT 70
Db 142 CTCTCAGTCTCTCAAAGCAAGGAAGAGACCGGTGCTGGAGACCATGGCAAGAAT 201
QY 71 CCTCCAGGAAATGTGAAGACTGTCACTTCTTAATGAGAACTTTTAAATCCAGAAA 130
Db 202 CCTCCAGGAAATGTGAAGGCTGTCACTTCTTAATGAGAACTTTCTGAAATCTAAGA 261
QY 131 ATATGTAATCACTTAAGATTGTGGACTGTGTGTTGTTATCTCTGGCCCTACTTAAT 190
Db 262 ATATGTAATCACTGAAGATTGTGGACTAGTG-TTGGTATCTCTGGCCCTTAACTTA 320
QY 191 GTCCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGCCCTATGACATG 250
Db 321 GTCCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTATCCAAGAAACCTATGACATG 380
QY 251 GAGCACACTTCTACAGCAATGGGAGAGAAAGAGATTTACATGGAATTTGATCCTGTG 310
Db 381 GAGCACACTTCTACAGCAATGGGAGAGAAAGAGATTTACATGGAATTTGATCCTGATA 440
QY 311 ACCAGAACTGAATATTACAGAAAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 370
Db 441 ACCAGAACTGAATATTACAGAAAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 499
QY 371 TTTAAAAACGATACATGCGCATCTACTCTGTTGGGTCTTCAAAAATGTTTATCAAAACT 430
Db 500 TTTAAAAATGGATACATGCGCATCTACTCTGTTGGGTCTTCAAAAATGCTTTTATTAAC 559
QY 431 CAGATTAAAGTGATTTCTGAAATTTTGAACCAAGAGAAATAGATGAGAAATGAAGAA 490
Db 560 CAAATCAAAGTGATTTCTGAAATTTTGAACCAAGAGAAATAGATGAGAAATGAAGAA 619
QY 491 ATTACCACAACTTCTTTGAACAGTCACTGATTTGGTCCCGACAGAAAAGCCCTATTGAA 550
Db 620 ATTACTACAACTTTC-TTGAACAGTCACTGATG-TTGGGTCTCCCGAGAAAAGCCCTAT--GA 675
QY 551 AACCGAGATTTCTTAAAAATTTCCAAAATTTCTG 583
Db 676 AACAGAGACTTCTGAAAATCTAAAATCTGCAG 708
```

RESULT 38

BF123957

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB317710 741 bp mRNA linear EST 04-MAR-2003  
AGENCOURT 12263597 NIH MGC 165 Mus musculus cDNA clone  
IMAGE:30281869 5', mRNA sequence.

CB317710.1 GI:28841945

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 741)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Leslie L. Heckert

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDCM41 row: m column: 14

High quality sequence stop: 422.

Location/Qualifiers

1..741

```
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30281869"
/tissue_type="primary cultures of Sertoli cells"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_165"
/notes="Organ: testis; Vector: pDNR-LIB; Site 1: Sf11
(ggcatatggcc); Site 2: Sf11 (ggcgcctcgcc); 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.6-3.5 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 35.0%; Score 419; DB 14; Length 741;
Best Local Similarity 88.7%; Pred. No. 2.4e-70;
Matches 489; Conservative 0; Mismatches 55; Indels 7; Gaps 3;

QY 11 CTCTCAGTCTCTCAAGCAAGCAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 70
DB 40 CTCTCAGTCTCTCAAGCAAGCAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 99
QY 71 CTCTCAGAGAAATGTGAAGTGTGACATCTTAATGCGAGAGCTTTAAATCCAGAAA 130
DB 100 CTCTCAGAGAAATGTGAGGGCTGTGACATCTTAATGCGAGAGCTTTAAATCCAGAAA 159
QY 131 ATATGTAATCACTTAAGATTTGTGAGCTGTGTTGGTATCTGCGCCCTAACTTAAT 190
DB 160 ATATGTAATCACTGAAGATTTGTGAGCTGTGTTGGTATCTGCGCCCTAACTTAAT 219
QY 191 GTCTCTTTTGGGGAGCAAGCACTTCTGCGCGAGGTATCCCAAAAAGCCTATGACATG 250
DB 220 GTCTCTTTTGGGGAGCAAGCACTTCTGCGCGAGGTATCCCAAAAAGCCTATGACATG 279
QY 251 GAGCACATTTTACAGCAATGAGAGAGAGAAATTTACATGGAATTTGATCTGCTG 310
DB 280 GAGCACATTTTACAGCAATGAGAGAGAGAAATTTACATGGAATTTGATCTGCTG 339
QY 311 ACCAGAACTGAATATTTACAGAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 370
DB 340 ACCAGAACTGAATATTTACAGAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 399
QY 371 TTTTAAAAACGGATACACTGGCATCTACTCTGCTGGTCTTCAAAAATGTTTTTATCAAAACT 430
DB 400 TTTTAAAAACGGATACACTGGCATCTACTCTGCTGGTCTTCAAAAATGTTTTTATCAAAACT 459
QY 431 CAGATTAAGTGAATCTCTG-AAATTTCTGAAACAGAGAGAGAAATGATGAGATGAAGA 489
DB 460 CAAATCAAAGTGAATCTCTGAAATGTTCTGAAACAGATGAAGAAATGATGAGATGAAGA 519
QY 490 AATTAC--ACAACCTTTCTTGAACAGTCAAGTGAATTT--GGGTCCCGAGCAGAAAAGCC 543
DB 520 AATTACTTACACTTTTCTTGAACAGTCAAGTGAATTTGGGGTTCGCGCAAAAAGCC 579
QY 544 TATTGAAAAC 554
DB 580 TATTGAAAAC 590

RESULT 39
BX638501/c
LOCUS
DEFINITION BX638501 pBluescript Lion Mus musculus cDNA clone EST 12-AUG-2003
3', mRNA sequence.
ACCSSION BX638501
VERSION BX638501.1 GI:33618376
KEYWORDS
SOURCE Mus musculus (house mouse)
```

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 521)  
REFERENCE  
AUTHORS Henrich, J., Hermanns, J., Kranz, H., Loebbert, R., Schluter, T.,  
Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
TITLE Mouse ArrayTAG cDNA (LION)  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; LiONp462G08420.  
RZPDLIB;  
Mouse ArrayTAG cDNA (LION)  
http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=4  
62 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
RP: CAGGAACACGTATGAC.  
FEATURES  
Location/Qualifiers  
1..521  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="LIONp462G08420"  
/lab\_host="DH10B"  
/clone\_lib="pBluescript Lion"  
ORIGIN  
Query Match 34.2%; Score 409.4; DB 13; Length 521;  
Best Local Similarity 87.9%; Pred. No. 1.9e-68;  
Matches 457; Conservative 0; Mismatches 59; Indels 4; Gaps 1;  
QY 668 GAAGATCTTCACCTTCTCTGCCAACGAAAAAGGATTGAAACAAATGACAGTGGTG 727  
DB 520 GAAGATCTTCACCTTCTCTGCCAACGAAAAAGGATTGAAACAAATGACAGTGGTG 461  
QY 728 GTCCCTCAAGTGAAGTAGAAGACCCGTCAGCCAGCAACGACGAGTGAAGAACTT 787  
DB 460 GTCCCGCAAGTGAAGTGGAGAGACCCGCCACACAGCAAGCAGGAGGAGACCTT 401  
QY 788 CCAATAAAGTACATATCTGAAATGGAATGAAATTTGATCCCATGCTGGATGAGAGAGT 847  
DB 400 CCTATAAATGACTATCTGAAATGGAATGGAATTTGACCCATGCTGGATGAGAGAGT 341  
QY 848 TATTGTTGATTTTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTA 907  
DB 340 TACTGTTGATTTTACTGTCGTCGAGGCAACCGTTACTGCGCGCGGTCTGTGAACCTTTA 281  
QY 908 CTAGGCTACTACCCATATCCATCTACTCTACAGGAGGAGAGTCACTCTCGTGTGTCATC 967  
DB 280 CTAGGCTACTACCCATATCCCTTACTCTCAAGGAGGTGAGTCACTCTCTGTGTGTCATC 221  
QY 968 ATGCTTTGTAACCTGTTGGGTGGCGCGCATGCTGGGAGGGGTCTAATAGGAGGTTTGAGCT 1027  
DB 220 ATGCTTTGCAACTGTGTGGGTGGCGCGCATGCTTTGGGAGAGTCTAATAGGAGATTTGAGTT 161  
QY 1028 CAAATGCTTAAACTGCT---GGCAACATATAATAATGATGATGCTATTCATGATTTCT 1083  
DB 160 CCNCGCTTAACTTCTGTTAGCCATATATATAATGATGCTATTCATGATTTCT 101  
QY 1084 GCCTATGAGCACTTGGCCCTGTAGCCAGCTCTCCAGAAATTAATCTGTAGTAATTCCT 1143  
DB 100 GCCTATGAGCAATTTGCTCCAGAGCTCTATCTTCAGAAATTAATCTGTAGTAATTCCT 41  
QY 1144 CTCTTCACTGTTCTAATAAACTTCTTACATTTATCACCAAAA 1183

```

Db      40  CTCCTCATGTTCTAATAAACTTCTACATCATCATCAAAAA 1
RESULT 40
AI039039/c      399 bp mRNA linear EST 30-JUN-1998
LOCUS      ox30b10.sl Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:1657819 3', similar to SW:CHM1_BOVIN_P17404 CHONDROMODULIN-I
PRECUSOR i, mRNA sequence.
AI039039
ACCESSION AI039039.1 GI:3278233
VERSION Homo sapiens (human)
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amerham
High quality sequence stop: 389.
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1657819"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'-
TGTACCAACTGTAAGTGGAGCGCGCTTAATTTTTTTTTTTTTT 3']".
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match      32.2%; Score 385; DB 9; Length 399;
Best Local Similarity 99.7%; Pred. No. 9.5e-64;
Matches 396; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      780  AAGAACTTCCAAATAAATGACTATACTGAAATGGAATGGAATTTGATCCCATGCTGGATG 839
Db      399  AAGAACTTCCAAATAAATGACTATACTGAAATGGAATGGAATTTGATCCCATGCTGGATG 340
QY      840  AGAGAGTTATTGTTGATTACTGCGGTGAGGCAACCGCTATTGCGCGCGCTGCTGTG 899
Db      339  AGAGAGTTATTGTTGATTACTGCGGTGAGGCAACCGCTATTGCGCGCGCTGCTGTG 280
QY      900  AACCTTTACTAGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTC 959
Db      279  AACCTTTACTAGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTC 220
QY      960  GTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGGTCTAATAGAGG 1019
Db      219  GTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGGTCTAATAGAGG 160
QY      1020  TTGAGCTCAATGCTTAACTGCTGGCAACATATAATAATGATGCTATTCAATGAAT 1079
Db      159  TTGAGCTCAATGCTTAACTGCTGGCAACATATAATAATGATGCTATTCAATGAAT 100
QY      1080  TTCTGCTATAGGCAATCGGCCCTTGAGGAGCTCTCCAGAAATTACTTGTAGGTAAT 1139

```

```

Db      99  TTCT-CCTATGAGCATCTGGCCCTGGTAGCGAGCTCTCCAGAAATTACTTGTAGTAAT 41
QY      1140  TCCTCTCTTCATGTTCTTAATAAACTTCTACATTATCA 1176
Db      40  TCCTCTCTTCATGTTCTTAATAAACTTCTACATTATCA 4
RESULT 41
CB202990      1224 bp mRNA linear EST 05-FEB-2003
LOCUS      AGENCOURT 11290775 NIH_MGC_135 Mus musculus cDNA clone
DEFINITION IMAGE:30139809 5', mRNA sequence.
CB202990
ACCESSION CB202990.1 GI:28238967
VERSION Mus musculus (house mouse)
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0044 row: n column: 10
High quality sequence start: 86
High quality sequence stop: 334.
FEATURES             Location/Qualifiers
     source          1..1224
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30139809"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_135"
                     /note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5 13.5,
14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'GACTAGTTCTAGATCGGCGGCC(T)3' Tissue contributed
by, David Rowe. Library constructed by ResGen, Invitrogen
Corp."
ORIGIN
Query Match      31.7%; Score 379.8; DB 14; Length 1224;
Best Local Similarity 85.0%; Pred. No. 6.5e-63;
Matches 472; Conservative 0; Mismatches 77; Indels 6; Gaps 4;

QY      12  TCTCAGTCTCTCAAGCA-AGGAAGAGTACTGTGCTGAGACCATGGCAAGAAT 70
Db      64  TCTCCGGATCTCAAGCAGGGGAAGAGCATCGTGTCTGGAGACCATGGCAAGAAT 123
QY      71  CCTCCAGAGAAATGTTGAAGACTGTACATTCTAAATGCAGAGCTTTTAAATCCAGAAA 130
Db      124  CCTCCAGAGAACTGTGAGGGCTGTACATTCTAAATGCAGAGCTCTGAAATCTAAGAAG 183
QY      131  ATATGTAATCATCTTAAGATTGTTGGACTGGTGTGTTGGTATCTCTGGCCCTAACTTAAT 190
Db      184  ATATGTAATCATCTGAAGATTGTTGGACTAGTGTGTTGGTATCTCTGGCCCTAACTTAAT 243
QY      191  GTCTCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAAGCCCTATGACATG 250

```

Best Local Similarity 83.3%; Pred. No. 1.3e-62;		Matches 539; Conservative 0; Mismatches 94; Indels 14; Gaps 9;	
QY	48	GCTCAGAGACCATGGCAAGAAATCCTCCAGAGAATTGTGAAGACTGTGCACATTTCTAAATG	107
Db	1	GCTGGGAGACCATGGCAAGAAATCCTCCAGAGAATTGTGAGGCTGTGCACATTTCTAAATG	60
QY	108	CAGAAGCTTTTAAATCCAGAAATATGTAATCACTTAAAGTTTGTGAGCTGGTGT-166	
Db	61	CAGAAGCTCTGAAATCTAAGAAGATATGTAATCACTAAGATTTTGTGAGCTAGATGTTA	120
QY	167	-GGTATCTGCGCCCTAACTCTAAATTTGTCCTGTTTGGGGGAGCAAGCACTTCTGCGCGA	225
Db	121	GGATATCTGCGCTTAACTCTAAATTTGTCCTGTTTGGGGGAGCAAGCACTTCTGCGCGA	179
QY	226	GGTACCCAAAAGGCTATGACATGGAGCACA-CT-TTCTACAGCAATGGAGAGAGAAGA	284
Db	180	GGTATCCAGAAACCTATGACATGGAGCACA-CTTCTACAGCAACGGGAGAGAAGA	239
QY	285	AGATTTCATGGAAATTGATCTCTGACGAGCACTTAAATATTCAGAACGCAATGGCA	344
Db	240	AGATTTCATGGAAATTGATCTCTGACGAGCACTTAAATATTCAGAACGCAATGGCA	299
QY	345	CTGATGAACATTGGAAGTCCATGACTTTAAATGGATACACTGGCACTTCTT-404	
Db	300	CTGATGAACATTGGAAGTCCATGACTTTAAATGGATACACTGGCACTTCTT-404	358
QY	405	GTCTTCAAAATGTTTATCAAACTCAGATTAAGTGAATTCCTGAAATTTCTGAACCG	464
Db	359	GTCTTCAAAATGTTTATCAAACTCAGATTAAGTGAATTCCTGAAATTTCTGAACCG	418
QY	465	AGAGAGAA-ATGATGAGAAATGAAGAA---TTACCAACAATTTCTTGAACAGTCAGTG	520
Db	419	AGAGAGAACTAGATGAGAAATGAAGAACTTACTACCACTTCTTGAACAGTCAGTG	478
QY	521	ATTGGGTCCTCCAGCAAGAAAGCTATTGAAACCGAGATTTCTTAAAAATTTCCAAAT	580
Db	479	ATTGGGTCCTCCAGCAAGAAAGCTATTGAAACCGAGATTTCTTAAAAATTTCCAAAT	538
QY	581	CTG---GAGATTGTGATTAACGTGACATGTA-TTGGATCAATCCCACTCTATATCAGT	636
Db	539	CTGGCAGATATGCGATACTGTGACCATGTACTTGGATCAATCACAATCTAATAGCAG	598
QY	637	TTCTGAGTTACAAGACTTTGAGGAGGAGGAGAGATCTTCACTTTC	683
Db	599	TATCGAATTACAGGAC-TTGACGAGGAGCGAGAAATTCATCATTC	644
RESULT 43			
AI146280/c			
LOCUS			
DEFINITION			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 825 Std Error: 0.00			
Seq primer: -40m13 fwd. ET from Amersham			
High quality sequence stop: 450.			
AI146280 460 bp mRNA linear EST 26-OCT-1998			
qp87b10.x1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone			
IMAGE:1707067 3, similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I			
PRECUSOR ; mRNA sequence.			
AI146280			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
This clone is available royalty-free through LNL; contact the			

FEATURES	SOURCE
----------	--------

Location/Qualifiers  
1. .460

```

1. 460
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1707067"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBH19W"
/note="Organ: heart; Vector: pTV773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAACTCGAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTV73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBH19W"

```

## ORIGIN

Query Match 31.4%; Score 376.4; DB 9; Length 460;  
Best Local Similarity 99.7%; Pred. No. 4e-62;  
Matches 377; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

	QY		801	A T A C T G A A A T G G A A T A G A A T T T G A T C C C A T G C T G G A T G A G A G A G G T A T T T G T G T A T T T	860
	D b		378	A G A C T G A A A A T G G A A T A G A A T T T G A I C C C A T G C T G G A T G A G A G A G G T A T T T G T G T A T T T	319
	QY		861	A C T G C C G T C G A G C A A C C G C T A T T G C C G C G C T G T G T G A A C T T T A C T A G G C T A C T P A C C	920
	D b		318	A C T G C C G T C G A G C A A C C G C T A T T G C C G C G C T C T G T G A A C T T T A C T A G G C T A C T P A C C	259
	QY		921	C A T A T C C A T A C T C T A C C A A G A G A G A C G A G T C A T C T G T C G T G C A T C A T G C C T T G T A A C T	980
	D b		258	C A T A T C C A T A C T G T A C C A A G A G A G A C G A G T C A T C T G T C G T G C A T C A T G C C T T G T A A C T	199
	QY		981	G G T G G T G G C C C G C A T G C T G G G A G G G T C T A A T A G G A G G T T T G A G C T C A A A T G C T T A A A C	1040
	D b		198	G G T G G T G G C C C G C A T G C T G G G A G G G T C T A A T A G A G A G G T T T G A G C T C A A A T G C T T A A A C	139
	QY		1041	T G C T G C R A C A T A T A N T A A T G C A T G C T A T T C A A T G A A T T T C T G C C H A T G A G C A C T G G	1100
	D b		138	T G C T G G C A A C A T A N T A A A T G C A T G C T A T T C A A T G A A T T T C T G C C T A T G A G C A C T G G	79
	QY		1101	C G C C T G G T A G C C A G C T C T C C A A A T T A C T T T G A G G T A A T T C C F C T C T T C A T G T C T T A N T A	1160
	d b		78	C C C C G G T A G C C A G C T C T C C A A A T T A C T T T A G G T A A T T C C T C T C T T A T G T T C T A N T A	19
	QY		1161	R A C T T C T A C A T T A T C A C C	1178
	D b		18	R A C T T C T A C A T T A T C A C C	1

RESULT 44

Accession	LOCUS	Size	Source	Library	EST
AV593196/c	AV593196	468 bp	Bos taurus	linear	EST 27-NOV-2001
	DEFINITION		Bos taurus cartilage fetus	cDNA clone	
			ElcA006A12 3', mRNA sequence.		

ACCESSION AV593196  
VERSION AV593196.1  
KEYWORDS EST. GI:9708353

SOURCE ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.  
1 (bases 1 to 468)  
Takasuga, A., Hirotsune, S., Itoh, R., Jitchazono, A., Suzuki, H., Aso, H.  
AUTHORS  
REFERENCE  
Y. Sugimoto.

## TITL E

Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs  
Nucleic Acids Res. 29 (22), E108 (2001)  
21570554  
11713328  
Contact: Yoshikazu Sugimoto

Contact: Ioshikazu Sugimoto  
 Animal Genetics Division  
 Shirakawa Institute of Animal G  
 Odakura, Nishigo, Nishi-shiraka  
 Tel: 81-248-25-5641  
 Fax: 81-248-25-5725  
 Email: kazuugi@cocoa.ocn.ne.jp

Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library.

	source
<hr/>	
FEATURES	

```

1. 468
   'organism="Bos taurus"
   /mol_type="mRNA"
   /db_xref="taxon:9913"
   /clones="ElCA006A12"
   /tissue_type="cartilage"
   /dev_stage="fetufus"
   /lab_host="DH10B"
   /clone_lib="Bos taurus c
   /note=vector: pZLL; Sit
   was deleted from a Nori

```

## ORIGIN

Query Match	31.0%;	Score 371.6;	DB 9;	Length 468;
Best Local Similarity	90.2%;	Pred. No. 3.3e-61;		
Matches 422;	Conservative 0;	Mismatches 39;	Indels 7;	Gaps 2;
QY	693	AAAAAAGGGATTGAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGA	752	
Db	468	AAAAAAGGCATTGAACAAAAACGACGCGTGGGTGGTCCCTCAAGTGAAGTGGAGAAGA	409	
QY	753	CCCGTCCAGCCGACACA---AGCAAGTGAAGGAACCTTCCAATAAATGACTATATCGAAA	809	
Db	408	CCCGTCACACACAGACAGCAGACAGTAGAGGAAGACTCCCAATCAATGACTATCTGAAA	349	
QY	810	ATGGAATAGAATTGATCCATCCATGCTGGATGAGAGAGGTTATTGTTGTTATTACTGCCGTC	869	
Db	348	ATGGAATAGAATTCGACTCCATGTTGGATGAGAGAGGTTACTGTGTTATTACTGCCGTC	289	
QY	870	GAGGCAACCGCTATTGCCCGCGGCTGTGTGAACCTTTACTAGGCTACTCCCATATCCAT	929	
Db	288	GAGGCAACCGCTACTGTGCCCGGCTGTGTGAACCTTTACTAGGTTACTACCCGTATCCAT	229	
QY	930	ACTGCTTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCTTGACTGCTGGGTGG	989	
Db	228	ATTGCTACCAAGAGGGCGGGTTATCTGTGCTGTCATCATGCTTGCACCTGGGTGGTGG	169	
QY	990	CCCGCATGTGGGAGGGGTCTAATAGGAGGTTTGAGCTCAAAATGCTTAAACTGCTGG---	1046	
Db	168	CCCGCATGCTAGGAGGGTTTAAATAGCAGTCTGAGTCGAGTGCTTAAACTTCTGGCAT	109	
QY	1047	CAACATATATAAATGATGCTATTCAATGAATTTCTGCCCTATGAGGCATCTGGGCCCT	1105	
Db	108	CCAAATATAACAAATGCATGCTATTCAATGAATTTCTGCCCTATGAGGCATTTGGTCTCT	49	
QY	1106	GGTAGCCAGCTCTCCAGAAATTACTGTGAGGTAAATTCCTCTCTTCATGT	1153	
Db	48	GGTAGCCAAATATCCGGAATTCGTTGTAGTAAATTCCTCTCTTCATGT	1	

## RESULT 45

CF112025	CF112025	433 bp	mRNA	linear	EST 23-JUL-2003
LOCUS	Shulzomic05276				
DEFINITION	Rat lung airway and parenchyma CDNA libraries				
	Rattus norvegicus CDNA clone CA2448 5', mRNA sequence.				
	CF112025				
ACCESSION	CF112025.1				
VERSION	GI:33169492				

KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM  
REFERENCE Rattus norvegicus (Norway rat)  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 433)  
Shultz M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,  
Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,  
Plopper, C.G. and Buckpitt, A.R.  
Gene Expression Analysis in Response to Lung Toxicants: I.  
Sequencing and Microarray Development  
Unpublished (2003)  
Contact: Shultz MA  
Dept. of Molecular Biosciences, School of Veterinary Medicine  
University of California, Davis  
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA  
Tel: 530 752 0793  
Fax: 530 752 4698  
Email: mashultz@ucdavis.edu  
Average Phred score is 20 or better. All poor quality data (Phred <  
20) and vector/linker sequence has been removed.  
High quality sequence stop: 433.  
Location/Qualifiers  
1. .433  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="CA2448"  
/sex="male"  
/tissue type="airway or parenchyma"  
/dev\_stage="adult"  
/clone\_lib="Rat lung airway and parenchyma cDNA libraries"  
/note="Organ: lung; Vector: pGEM-11zf(-); Site\_1: Eco RI;  
Site 2: Not I; mRNA was isolated from microdissected rat  
lung airways and parenchyma tissues."  
FEATURES  
source  
Query Match 30.8%; Score 369; DB 14; Length 433;  
Best Local Similarity 90.8%; Pred. No. 1.1e-60;  
Matches 393; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 297 AAATTGATCTGTGACACAGAACTGAAATATTGAGAGCGGAATGGCACTGATGAACAT 356  
DB 1 AAATTGATCCATAACACAGACAGAAATATTGAGAGTGGAAATGGCACCGATGAACAT 60  
QY 357 TGAAGTGCACGACTTTTAAACACGATACACTGGCATCTACTTCGTGGTCTTCAAAAT 416  
DB 61 TGAAGTCCATGACTTTTAAACACGATACACTGGCATCTACTTTGTAGTCTTCAAAAT 120  
QY 417 GTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGACACAGAGAGGAATAG 476  
DB 121 GCTTTATTAACCTCAATCAAGTAGTAACTCTGAAATTTCTGACACAGAGGAATAG 180  
QY 477 ATGAGATGAAGAAATTAACCAACTTCTTTGAACAGTCAGTCAGTATTTGGTCCACGAG 536  
DB 181 ATGAGATGAGAAATTAACACGTTCTTTGAACAGTCAGTCAGTATTTGGTCTTCGAG 240  
QY 537 AAAAGCCTATTGAAACCGAGATTTCTTAAAAATTCCAAATTTCTGGAGTTTGTGATA 596  
DB 241 AAAAGCCTATTGAAACAGAGACTTCTGAAATTTCTGAAATTTCTGGAGATTTGCGACA 300  
QY 597 ACGTGACCATGATTGGATCAATCCACTCTAATATCAGTTTCTGAGTTTCAAGACTTTG 656  
DB 301 ATGTGACTATGTACTGGATCAATCCACTCTAATATGAGTTTCTGAGTTTCAAGACTTTG 360  
QY 657 AGGAGGAGGAGAGATCTTCACTTTCTGCGCAAGCAAAAAAGGGATTGACAAAATG 716  
DB 361 AGGAGGATGGTGAAGATCTTCACTTTCTACACGCAAAAAAGGGATTGACCAAGATG 420  
QY 717 AACAGTGGGTGGT 729  
| | | | |

Db 421 AGCAATGGTGGT 433  
RESULT 46  
LOCUS A1112003/c  
DEFINITION UI-R-Y0-mp-b-11-0-UI.s1 UI-R-Y0 Rattus norvegicus cDNA clone  
VERSION A1112003  
KEYWORDS UI-R-Y0-mp-b-11-0-UI 3', mRNA sequence.  
SOURCE A1112003.1 GI:3511952  
ORGANISM EST.  
Rattus norvegicus (Norway rat)  
REFERENCE Rattus norvegicus  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
JOURNAL Rattus.  
MEDLINE 1 (bases 1 to 487)  
PUBMED Bonaldo, M.F., Lennon, G. and Soares, M.B.  
COMMENT Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
The sequence tag present in the cDNA between the NotI site and the  
oligo-dT track served to identify it as a clone from the normalized  
adult Eye library. cDNA library Preparation: M. Fatima Bonaldo,  
Ph.D. Clone distribution: clones will be available through Research  
Genetics  
Seq primer: M13 Forward.  
Location/Qualifiers  
1. .487  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-Y0-mp-b-11-0-UI"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-Y0  
library is a subtracted library derived from an  
individually-tagged normalized whole-eye (minus the lens)  
pool of all previous libraries (UI-R-A0, UI-R-A1, UI-R-E0,  
UI-R-E1, UI-R-C0, and UI-R-C1). The tag is a string of  
3-5 nucleotides present between the Not I site and the  
oligo-dT track which allows identification of the library  
of origin of a clone within the mixture. The subtracted  
library (UI-R-Y0) was constructed as follows: PCR  
amplified cDNA inserts from previous library clones from  
which 3' ESTs had been derived were used as a driver in a  
hybridization with the normalized whole-eye library in  
the form of single-stranded circles. The remaining  
single-stranded circles (subtracted library) was purified  
by hydroxyapatite column chromatography, converted to  
double-stranded circles and electroporated into DH10B  
bacteria (Life Technologies) to generate the UI-R-Y0  
library. This procedure has been previously described  
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,  
1996)"  
ORIGIN  
Query Match 30.6%; Score 366.8; DB 9; Length 487;  
Best Local Similarity 86.9%; Pred. No. 2.7e-60;  
Matches 431; Conservative 0; Mismatches 52; Indels 13; Gaps 2;



```

QY 704 ATTGAACAAATGAACTAGTGGTGGTCCCTCAAGTGAAAGTGAAGAGACCCGTCACGCC 763
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 ATTGACAGAAATGAGCAATGGTGGTCCCAACAAGTGAAGTGAAGAGACCCGCGACC 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 764 AGACAAGCAAGTGAAGGAACCTTCCAATAAATGACTATCTACTGAAATGGAATGAAATTT 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 AGACAAGCAAGCGAGGAGACCTTCTCTGTTAATGACTATCTACTGAAATGGAATGAAATTT 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 824 GATCCCATGCTGGATGAGAGAGTTATTTGTTATTTACTGCGGTGAGGCAACCGCTAT 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GATCCCATGCTGGATGAGAGAGTTACTGTTGTTATTTACTGCGGTGAGGCAACCGCTAC 308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 884 TGCGCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTACTGTTACCAAGGA 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TGCGCGAGGCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTACTGTTACCAAGGA 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 944 GGACGAGTCTATCTGCTGTCATCATGCTCTGTAAGTGGTGGTGGCGCGCATGCTGGG 1003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 GGTGAGTCTATCTGCTGTCATCATGCTCTGTAAGTGGTGGTGGCGCGCATGCTGGG 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1004 AGGCTTAATAGAGGTTTGGAGTCAAAATGCTTAAACTGCT---GGCAACATATAATAA 1059
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 187 AGAGTCTAATAGGAAGTTTGAGTCCAAATGCTTAACTTTTGTAGCCAAATATAATTA 128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1060 ATGCATGCTATTCAATGAATTTCTGCTTAAGGAGCATCTGGCCCTGTPAGCAGCTCTC 1119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127 ATGCATGCTACTCCATGAATTTCT-----GCATTTGCCCTCAAGTAGCTATCCTC 77
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1120 CAGAAATTTAGTATCTCTCTCTCTCANGTTCTTAATAAATCTTACATTTATCA 1179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 CAGAAATTTAGTATCTCTCTCTCTCANGTTCTTAATAAATCTTACATTTATCA 17
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1180 AAAAAAAAAAAAAA 1195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 16 AAAAAAAAAAAAAA 1
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 47
BF523627
LOCUS
DEFINITION
504 bp mRNA linear EST 11-DEC-2000
UI-R-C1-1c-e-10-0-UI.r1 UI-R-C1 Rattus norvegicus cdna clone
UI-R-C1-1c-e-10-0-UI 5', mRNA sequence.
ACCESSION
BF523627
VERSION
BF523627.1 GI:11631594
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 504)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
97044477
MEDLINE
889548
PUBMED
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
cdna Library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1792597
Seq primer: M13 Forward.
FEATURES
Location/Qualifiers
1..504
/organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C1-1c-e-10-0-UI"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C1"
/notes="Vector: p7713D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C1
library is a subtracted library derived from the UI-R-C0
library, which is a subtracted library derived from the
UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library
consisted of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The UI-R-E1 library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-C1) was constructed as follows: PCR
amplified cDNA inserts from UI-R-C0 clones from which 3'
ESTs had been derived was used as a driver in a
hybridization with the UI-R-C0 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C1
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)."
ORIGIN
Query Match
Best Local Similarity 30.6%; Score 366.2; DB 10; Length 504;
Matches 444; Conservative 0; Mismatches 53; Indels 16; Gaps 3;
QY 668 GAAGATCTTCACTTTCTTCGCCAACGAAAAAAGGATTGAACAAAATGAACAGTGGTG 727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4 GAGGATCTTCACTTTCTTACCAGCGAAAAAAGGATTGACAGATGAGCAATGGGTG 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 728 GTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGCAAGCAAGCAAACTT 787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GTCCCAAGTGAAGTAGAAGACCCGTCACGCCAGCAAGCAAACTT 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 CCAATAATGACTATCTGAAAAATGGAATAGAAATTCATCCATGCTGGATGAGAGGT 847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 122 CCTGTTAATGACTATCTGAAAAATGGAATAGAAATTCATCCATGCTGGATGAGAGGT 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 TATGTTGTTATTTACTGCGTCGAGGCAACCGTATTGCGCGCCGCTGTGAACCTTTA 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 182 TACTGTTGTTATTTACTGCGTCGAGGCAACCGCTACTGCGCGAGGTCTGTGAACCTTTA 241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 CTAGGCTTACTACCCATATCCATCTGCTACCAAGGAGGACGATCTCTCTGTCATC 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 CTAGGCTTACTACCCATATCCATCTGCTACCAAGGAGGTCGATCTCTCTGTCATC 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 ATGCTTTGTTAACTGGTGGTGGCGCCCATGCTGGGAGGCTCTAATAGAGGTTTGGCT 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 302 ATGCTTTGCACTGGTGGTGGCGCCCATGCTGGGAGAGTCTAATAGGAGTTTGGATC 361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1028 CAAATGCTTAACTGCT----GGCAACATATAATAATGCTATTCATCAATTTCT 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 362 CAAATGCTTAACTTTTGTAGCCACATATAATTAATGATGCTACTCTCATGAAATTTCT 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1084 GCCTATGAGCATCTGGCCCTGGTGGAGCAGCTCTCCAGAAATTTCTGTAGTAAATTCCT 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 422 GCAT-----TTGCGCTCAAGTAGCCTATCTCCAGAAATTTATTGTAGGATATTCCT 471
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1144 CTCTTCATGTTCTAATAAACTTCTACATTTATCA 1176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

Library (UI-R-Cl) was constructed as follows: PCR amplified cDNA inserts from UI-R-CO clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-CO library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-Cl library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)." source

## ORIGIN

Query Match 30.0%; Score 358.8; DB 9; Length 492;  
Best Local Similarity 86.5%; Pred. No. 9.2e-59;  
Matches 436; Conservative 0; Mismatches 52; Indels 16; Gaps 3;  
  
QY 693 AAAAAAAAAAGGATTGAACAATAAGACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGA 752  
Db |||||  
QY 492 AAAAAAAAAAGGATTGACCAAGATGACCAATGGTGGTGGTCCCAAGTGAAGTGGAGAAGA 433  
Db |||||  
QY 753 CCGGTCACGCCAGACCAAGTGAAGAGACTTCCAAATGAATGACTATGCTGAAATG 812  
Db |||||  
QY 432 CCGCGCCAGCAGACAGACAA--GAGGAAGACCTTCCTGTTAATGACTATGCTGAAATG 375  
Db |||||  
QY 813 GAATAGAAATTTGATCCCATGCTGATGAGAGAGGTTATTGTTGTTATTTACTGCCGTCGAG 872  
Db |||||  
QY 374 GAATCGAATTTGATCCCATGCTGGATGAGAGGTTACTGTTGTTATTTACTGCCGTCGAG 315  
Db |||||  
QY 873 GCAACCGCTATGCGCGCGCTGTGTAACCTTTACTAGGCTACTACCATATCCATCTACT 932  
Db |||||  
QY 314 GCAACCGCTACTGCGCGAGGCTGTGTAACCTTTACTAGGCTACTACCATATCCCTACT 255  
QY 933 GCTACCAAGGAGGAGAGTCACTGTCGTGTCATGCTGCTGTAACCTGCTGGTGGGCC 992  
Db |||||  
QY 254 GCTACCAAGGAGGTCGATCATCTGTCGTGTCATGCTGCTGTAACCTGCTGGTGGGCC 195  
Db |||||  
QY 993 GCATGCTGGGAGGCTCTAATAGGAGGTTTGAAGTCAAACTGCTAACTGCT----GGCA 1048  
Db |||||  
QY 194 GCATGCTGGGAGAGTCTAATAGGAGGTTGAGTCCAAATGCTTAACCTTTGTTAGCCA 135  
QY 1049 ACATATAATTAATGATGCTATTCAATGAATTTCTGCTATGAGGCACTGCGCCCTGGT 1108  
Db |||||  
QY 134 ACATATAATTAATGATGCTATTCAATGAATTTCTGCTATGAGGCACTGCGCCCTGGT 85  
Db |||||  
QY 1109 AGCCAGCTCTCCAGAAATTTACTGTTAGTAAATTTCTCTCTCTCTCTCTCTCTCTCT 1168  
Db |||||  
QY 84 AGCTATCTCTCCAGAAATTTTGTAGGATATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 25  
Db |||||  
QY 1169 CATTATCAACCAAAAAAAAAAAAAA 1192  
Db |||||  
QY 24 CATCATCAAAAAAAAAAAAAAAAAA 1

RESULT 50  
BF679243  
LOCUS 602153429F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294374 5',  
DEFINITION mRNA sequence.  
ACCESSION BF679243  
KEYWORDS BF679243.1 GI:11953138  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 882)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: c9apbs-femail.nih.gov

Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1143 row: j column: 07  
High quality sequence stop: 436.  
Location/Qualifiers  
1..882  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294374"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 83"  
/note="Organ: prostate; Vector: pDNR-LiB (Clontech); Site 1: Sfil (ggcgctcgcc); Site 2: Sfil (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Query Match 29.9%; Score 358.4; DB 10; Length 882;  
Best Local Similarity 87.7%; Pred. No. 9.1e-59;  
Matches 556; Conservative 0; Mismatches 51; Indels 27; Gaps 14;  
  
QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAAAGTACTGTGCTCAGAGACCAT 60  
Db |||||  
QY 27 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAAAGTACTGTGCTCAGAGACCAT 86  
Db |||||  
QY 61 GGCAAGAATCTCCAGAGAATTTGGAAGTGTGCATCTTAATGAGAAGCTTTTAA 120  
Db |||||  
QY 87 GGC--AAGAATCTCCAGAGAATTTGGAAGTGTGCATCTTAATGAGAAGCTTTTAA 145  
Db |||||  
QY 121 ATCCAGAGAATATGTAATCAGTAAATTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGC 240  
Db |||||  
QY 146 ATCCAGAGAATATGTAATCAGTAAATTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGC 264  
Db |||||  
QY 181 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGC 299  
Db |||||  
QY 206 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGC 324  
Db |||||  
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAAGAAATTTACAT--GGAAG 359  
Db |||||  
QY 265 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAAGAAATTTACAT--GGAAG 384  
Db |||||  
QY 300 TTGATCTCTGTGACCAAGCAAGTAAATTTACAGAGCGGAATGGCACTGATGAACATTGG 415  
Db |||||  
QY 325 TTGATCTCTGTGACCAAGCAAGTAAATTTACAGAGCGGAATGGCACTGATGAACATTGG 444  
Db |||||  
QY 360 -AAGTGCAGCACTTTAAA--AAGCGATACACTGGCATCTACT--TCGTGGGTCTTCAAAA 468  
Db |||||  
QY 385 AAAGTGCAGCACTTTAAAACACCGGATACCACTGGCATCTACTCTCTGGGTCTTCAAAA 504  
Db |||||  
QY 416 ----TGTTTATCAAAACTCAGATT--AAAGTGAATTCCTGAATTT--TCTGAACCAAGAA 521  
Db |||||  
QY 445 CATGCTCTTATCACACACCTCAGATTACAGAGTGAATTCCTGAATTTAACTGAACCAAGAA 564  
Db |||||  
QY 469 GGAATA-----GATGAGAATGAAGAAATTCACCAACTTTCTTTGA--ACAGTCAAGTA 581  
Db |||||  
QY 505 AGGAACACTAGATGAAGACTGACAGAAACTTACCAACTTTCTTTTGACACAGTCAAGTA 620  
Db |||||  
QY 522 TTTGGGTCCAGCAGAAAAAGCCCTATTGAAACCGAGATTTTCTTAAAAATTTCAAAATTC 640  
Db |||||  
QY 565 TT--GGGTCCAGCAGAAAAAGGCTAT--GAAACGAGATTTTCTTAACATTCAGAAATTC 650  
Db |||||

Qy 582 TGGAG--ATTTCGTGATACGTGACCATGTATTGG 613  
||| |||  
Db 621 TGGCGAATGGTGACAAACCGTGACCCGTATTGG 654  
||| |||

Search completed: September 3, 2004, 09:20:16  
Job time : 2377 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 04:35:02 ; Search time 3269 Seconds  
(without alignments)  
15870.790 Million cell updates/sec

Title: US-10-063-730-115

Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtcctcc.....caaaaaaaaaaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1197	100.0	1197	6	AX092384	AX092384 Sequence
2	1197	100.0	1197	6	AX464390	AX464390 Sequence
3	1197	100.0	1197	6	AX697253	AX697253 Sequence
4	1197	100.0	1197	9	AY358706	AY358706 Homo sapi
5	1196	99.9	1228	6	BD269275	BD269275 33 human
6	1194	99.7	1206	6	AX203127	AX203127 Sequence
7	1187.2	99.2	1200	6	BD094021	BD094021 A novel p
8	1178	98.4	1178	6	BD228713	BD228713 Mammalian
9	1174	98.1	1380	6	AR338895	AR338895 Sequence
10	1173.4	98.0	1214	9	AF234259	AF234259 Homo sapi
11	1173	98.0	1309	9	AF191770	AF191770 Homo sapi
12	1170	97.7	1184	9	AF291656	AF291656 Homo sapi
13	1138.4	95.1	1140	9	AB055421	AB055421 Homo sapi
14	1116	93.2	1428	6	AR338895	AR338895 Sequence
15	1043.6	87.2	1175	4	AB059407	AB059407 Equus cab
16	994.6	83.1	1338	10	BC049944	BC049944 Mus muscu
17	994.6	83.1	1338	10	AF291655	AF291655 Mus muscu
18	982.4	82.1	1202	10	AF219993	AF219993 Mus muscu
19	979.2	81.8	1180	6	BD094022	BD094022 A novel p
20	975	81.5	1197	6	BD094023	BD094023 A novel p
21	964	80.5	1336	10	AF191769	AF191769 Rattus no
22	961.8	80.4	1153	10	BC006919	BC006919 Mus muscu
23	949.4	79.3	951	6	BD077398	BD077398 Novel pro
24	949.4	79.3	951	6	BD096254	BD096254 Novel pro
25	949.4	79.3	975	6	BD094032	BD094032 A novel p
26	943.6	78.8	1128	10	AB055422	AB055422 Mus muscu
27	932.6	77.9	1214	10	AF191768	AF191768 Mus muscu
28	926.2	77.4	1136	10	AB055423	AB055423 Rattus no
29	823	68.8	951	6	AX203129	AX203129 Sequence
30	823	68.8	975	6	BD094033	BD094033 A novel p
31	805.4	67.3	807	6	BD077399	BD077399 Novel pro
32	805.4	67.3	807	6	BD096255	BD096255 Novel pro
33	378.8	31.6	97189	9	HS47937	AL035608 Human DNA
34	377.2	31.5	20850	6	AX203136	AX203136 Sequence
35	352.4	29.4	432	6	BD094036	BD094036 A novel p
36	309	25.8	309	6	BD077400	BD077400 Novel pro
37	309	25.8	309	6	BD096256	BD096256 Novel pro
38	302	25.2	221461	10	AL691421	AL691421 Mouse DNA
39	259	21.6	286	6	AX884585	AX884585 Sequence
40	259	21.6	286	6	BD024195	BD024195 Sequence
41	259	21.6	289	6	BD204197	BD204197 5' EST and
42	247	20.6	248	6	BD228719	BD228719 Mammalian
43	184.4	15.4	240491	2	AC107412	AC107412 Rattus no
44	134	11.2	1347	10	BC045152	BC045152 Mus muscu
45	130.8	10.9	1449	10	MM043509	U43509 Mus musculu

ALIGNMENTS

RESULT 1  
AX092384  
LOCUS AX092384 1197 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 115 from Patent WO0116318.  
ACCESSION AX092384  
VERSION AX092384.1 GI:13444506  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,  
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and  
Wood,W.I.

Db	841	GAGAGGTTATTGTTGATTATTACTGCGCGTGCAGGCAACCGCTATTTCGCCGCGCGTCTGTGA	900
Qy	901	ACCTTTACTAGGCTACTGCCATATCCATCTCTACCAAGGAGGACGAGTTCATCTCTCG	960
Db	901	ACCTTTACTAGGCTACTGCCATATCCATCTCTACCAAGGAGGACGAGTTCATCTCTCG	960
Qy	961	TGTCATCATGCCCTTGTAACTGTTGGGTGGCCCGCATCTGGGGAGGGTCTTAATAGGAGGT	1020
Db	961	TGTCATCATGCCCTTGTAACTGTTGGGTGGCCCGCATCTGGGGAGGGTCTTAATAGGAGGT	1020
Qy	1021	TTGAGCTCAAACTGCTTAAACTGCTGGCAACATATATAAATGATCTATTCATGAATT	1080
Db	1021	TTGAGCTCAAACTGCTTAAACTGCTGGCAACATATATAAATGATCTATTCATGAATT	1080
Qy	1081	TCCTGCTATGAGGCATCTGGCCCTTGTTAGCCAGCTCTCCAGAAATTACTTGTAGTAAAT	1140
Db	1081	TCCTGCTATGAGGCATCTGGCCCTTGTTAGCCAGCTCTCCAGAAATTACTTGTAGTAAAT	1140
Qy	1141	CTCTCTTCATGTTCTAAATAAACTTCTACATTTATCACCAAAAAAATAAAAAA	1197
Db	1141	CTCTCTTCATGTTCTAAATAAACTTCTACATTTATCACCAAAAAAATAAAAAA	1197
RESULT 2			
AX464390		1197 bp	DNA
LOCUS	AX464390	Sequence 523 from Patent WO0140466.	linear
DEFINITION	AX464390		
ACCESSION	AX464390.1	GI:21899211	
VERSION			
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E.,	
		Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,	
		Sherwood, S., Smith, V., Stewart, T.A., Tamas, D., Watanabe, C.K.,	
		Wood, W.L. and Zhang, Z.	
TITLE		Secreted and transmembrane polypeptides and nucleic acids encoding	
		same	
JOURNAL		Patent: WO 0140466-A 523 07-JUN-2001;	
		Genentech Inc. (US)	
FEATURES		Location/Qualifiers	
source		1. .1197	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		100.0%;	Score 1197; DB 6; Length 1197;
Best Local Similarity		100.0%;	Fred. No. 4.2e-286;
Matches 1197; Conservative		0; Mismatches	0; Indels
		0; Gaps	0
Qy	1	CAGCAGTGGTCTCTCAGTCTCTCTCAAAGCAAGAAAGTACTGTGCTGAGACCAT	60
Db	1	CAGCAGTGGTCTCTCAGTCTCTCTCAAAGCAAGAAAGTACTGTGCTGAGACCAT	60
Qy	61	GGCAAGAATCCTCCAGAGATTGTGAAGCTGTCACTTCTAAATGCAGAGCTTTTAA	120
Db	61	GGCAAGAATCCTCCAGAGATTGTGAAGCTGTCACTTCTAAATGCAGAGCTTTTAA	120
Qy	121	ATCCAAGAAATATGTAATACATTTAAGATTGTGGACTGTGTTTGTGTATCTCGCCCT	180
Db	121	ATCCAAGAAATATGTAATACATTTAAGATTGTGGACTGTGTTTGTGTATCTCGCCCT	180
Qy	181	AACCTTAATTCTGTTTGGGGAGCAAGACTTCTGGCCGGAGGTACCCAAAAAGC	240
Db	181	AACCTTAATTCTGTTTGGGGAGCAAGACTTCTGGCCGGAGGTACCCAAAAAGC	240
Qy	241	CTATGACATGAGCACATTTCTACAGCAATGGAGAGAAAGATTTACATGGAAT	300
Db	241	CTATGACATGAGCACATTTCTACAGCAATGGAGAGAAAGATTTACATGGAAT	300

QY 301 TGATCCTGTGACGAGAACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTGGA 360  
Db 301 TGATCCTGTGACGAGAACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTGGA 360  
QY 361 AGTGACGAGCTTTAAACCGGATACACTGGGATCTACTTCCTGGGGCTTCACAAAATGTTT 420  
Db 361 AGTGACGAGCTTTAAACCGGATACACTGGGATCTACTTCCTGGGGCTTCACAAAATGTTT 420  
QY 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAAACAGTCAGTGATTTGGTCCCGAGCAAAA 480  
Db 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAAACAGTCAGTGATTTGGTCCCGAGCAAAA 480  
QY 481 GAATGAAGAAATACACAGAACTTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 540  
Db 481 GAATGAAGAAATACACAGAACTTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 540  
QY 541 GCCTATTGAAACCGAGATTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 600  
Db 541 GCCTATTGAAACCGAGATTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 600  
QY 601 GACCATGATTGGATCAATCCACACTAATATCAGTTTCCAAAATTCGGAGATTTGTGATAACGT 660  
Db 601 GACCATGATTGGATCAATCCACACTAATATCAGTTTCCAAAATTCGGAGATTTGTGATAACGT 660  
QY 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAAACA 720  
Db 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAAACA 720  
QY 721 GTGGTGTGCTCAAGTGAAGATGAGAAGACCCGTCACGCGCAGCAAGCAAGTGAGGA 780  
Db 721 GTGGTGTGCTCAAGTGAAGATGAGAAGACCCGTCACGCGCAGCAAGCAAGTGAGGA 780  
QY 781 AGAACTTCCAAATAAGTACTATCTGAAATGGAATGAAATTTGATCCCATGCTGGATGA 840  
Db 781 AGAACTTCCAAATAAGTACTATCTGAAATGGAATGAAATTTGATCCCATGCTGGATGA 840  
QY 841 GAGAGGTTATTGTTGATTTACTGCGCTCGAGGCAACCGCTATTGCGCGCGGCTCTGGA 900  
Db 841 GAGAGGTTATTGTTGATTTACTGCGCTCGAGGCAACCGCTATTGCGCGCGGCTCTGGA 900  
QY 901 ACCTTTACTAGGCTACTACCATATCCATATCTGCTACCAAGGAGGACGAGTCACTGTG 960  
Db 901 ACCTTTACTAGGCTACTACCATATCCATATCTGCTACCAAGGAGGACGAGTCACTGTG 960  
QY 961 TGTCAATCATGCTTGAATGCTGGTGGTGGCGGATGCTGGGAGGGTCTAATAGGAGGT 1020  
Db 961 TGTCAATCATGCTTGAATGCTGGTGGTGGCGGATGCTGGGAGGGTCTAATAGGAGGT 1020  
QY 1021 TTGAGCTCAAATGCTTAACTGCTGGCAACATATAATAAATGCATGCTATTCAATGAATT 1080  
Db 1021 TTGAGCTCAAATGCTTAACTGCTGGCAACATATAATAAATGCATGCTATTCAATGAATT 1080  
QY 1081 TCTGCTTATGAGGATCTGGCCCTGTAGCCAGCTCTCCAGAAATTAATGAGTAATT 1140  
Db 1081 TCTGCTTATGAGGATCTGGCCCTGTAGCCAGCTCTCCAGAAATTAATGAGTAATT 1140  
QY 1141 CCTCTCTCATGTTCTAATAAATCTTACATTTACCAACAAAAAAGGATTTGATGAGTAATT 1197  
Db 1141 CCTCTCTCATGTTCTAATAAATCTTACATTTACCAACAAAAAAGGATTTGATGAGTAATT 1197

RESULT 3  
AX697253  
LOCUS AX697253 1197 bp DNA linear PAT 02-APR-2003  
DEFINITION Sequence 321 from Patent W00078961.  
ACCESSION AX697253  
VERSION AX697253.1 GI:29498415  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L.,  
Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A.,  
Godowski, P.J., Gurney, A.L., Smith, V., Tumas, D., Wood, W.I.,  
Grimaldi, C.J., Hillan, K.J., Paoni, N.F., Roy, M.A. and Watanabe, C.K.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
the same  
JOURNAL Patent: WO 0078961-A 321 28-DEC-2000;  
Genentech Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..1197  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1197; DB 6; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 4.2e-286;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAGCAGTGGTCTCTCAGTCCCTCTCAAGCAAGAAAGTAGTACTGTGCTGAGAGACCAT 60  
Db 1 CAGCAGTGGTCTCTCAGTCCCTCTCAAGCAAGAAAGTAGTACTGTGCTGAGAGACCAT 60  
QY 61 GGCAGAGAAATCCTCCAGAGAAATGTAAGACTGTGAAAGCTGTACATTTCTAAATGCAGAGCTTTTAA 120  
Db 61 GGCAGAGAAATCCTCCAGAGAAATGTAAGACTGTGAAAGCTGTACATTTCTAAATGCAGAGCTTTTAA 120  
QY 121 ATCCAAAGAAATATGTAATCATTAAATTTGTGGACTGTGCTGCTGCTGGCCCT 180  
Db 121 ATCCAAAGAAATATGTAATCATTAAATTTGTGGACTGTGCTGCTGCTGGCCCT 180  
QY 181 AACTCTAATTTGCTGTTTTGGGGAGCAACACTTTCTGGCGGAGGTACCCAAAAAGC 240  
Db 181 AACTCTAATTTGCTGTTTTGGGGAGCAACACTTTCTGGCGGAGGTACCCAAAAAGC 240  
QY 241 CTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAGAAAGATTTCATGGAAT 300  
Db 241 CTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAGAAAGATTTCATGGAAT 300  
QY 301 TGATCCTGTGACAGAACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTGGA 360  
Db 301 TGATCCTGTGACAGAACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTGGA 360  
QY 361 AGTGACGAGCTTTAAACCGGATACACTGGGATCTACTTCCTGGGGCTTCACAAAATGTTT 420  
Db 361 AGTGACGAGCTTTAAACCGGATACACTGGGATCTACTTCCTGGGGCTTCACAAAATGTTT 420  
QY 421 TATCAAACTCAGATTAAAGTGATTCCTGAAATTTCTGAAACAGTCAGTGATTTGGTCCCGAGCAAAA 480  
Db 421 TATCAAACTCAGATTAAAGTGATTCCTGAAATTTCTGAAACAGTCAGTGATTTGGTCCCGAGCAAAA 480  
QY 481 GAATGAAGAAATACACAGAACTTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 540  
Db 481 GAATGAAGAAATACACAGAACTTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 540  
QY 541 GCCTATTGAAACCGAGATTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 600  
Db 541 GCCTATTGAAACCGAGATTTCTTGAACAGTCAGTGATTTGGTCCCGAGCAAAA 600  
QY 601 GACCATGATTGGATCAATCCACACTAATATCAGTTTCCAAAATTCGGAGATTTGTGATAACGT 660  
Db 601 GACCATGATTGGATCAATCCACACTAATATCAGTTTCCAAAATTCGGAGATTTGTGATAACGT 660  
QY 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAAACA 720  
Db 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAAACA 720  
QY 721 GTGGTGTGCTCAAGTGAAGATGAGAAGACCCGTCACGCGCAGCAAGCAAGTGAGGA 780  
Db 721 GTGGTGTGCTCAAGTGAAGATGAGAAGACCCGTCACGCGCAGCAAGCAAGTGAGGA 780  
QY 781 AGAACTTCCAAATAAGTACTATCTGAAATGGAATGAAATTTGATCCCATGCTGGATGA 840





QY	ACC	TTT	TACTAG	CTACTACT	CCCATATCC	CACTACTG	TACCAAGG	AGGACG	AGTCA	TCCTCG	960	
Db	ACC	TTT	TACTAG	CTACTACT	CCCATATCC	CACTACTG	TACCAAGG	AGGACG	AGTCA	TCCTCG	960	
QY	TGT	CAT	CATG	CTTTG	TAACTG	GTGGGTG	GGCCG	CATGCTG	GGGAGG	GTCTAATAG	AGGT 1020	
Db	TGT	CAT	CATG	CTTTG	TAACTG	GTGGGTG	GGCCG	CATGCTG	GGGAGG	GTCTAATAG	AGGT 1020	
QY	TTG	AGCT	CAAA	TGCTT	AAATG	CTGTCG	GCAACAT	ATATAA	TAATAT	GAATGAT	CTTCAATGA	ATT 1080
Db	TTG	AGCT	CAAA	TGCTT	AAATG	CTGTCG	GCAACAT	ATATAA	TAATAT	GAATGAT	CTTCAATGA	ATT 1080
QY	TCTG	CCCTAT	GAGG	CATCTG	GGCCCTG	TAGCG	AGCTCTC	CAGAA	TACTTCTG	TAGGTA	ATT 1140	
Db	TCTG	CCCTAT	GAGG	CATCTG	GGCCCTG	TAGCG	AGCTCTC	CAGAA	TACTTCTG	TAGGTA	ATT 1140	
QY	CCTCT	CTTCAT	GTCTA	ATAA	CTTCA	CAATAT	CACCA	AAAAA	AAAAA	AAAAA	AAAAA 1197	
Db	CCTCT	CTTCAT	GTCTA	ATAA	CTTCA	CAATAT	CACCA	AAAAA	AAAAA	AAAAA	AAAAA 1197	
RESULT 5	BD269275					1228 bp	DNA	linear	PAT 17-JUL-2003			
LOCUS	33 human secreted proteins.											
DEFINITION	BD269275											
ACCESSION	BD269275.1	GI:33079043										
VERSION	JP 2002534972-A/10.											
KEYWORDS	Homo sapiens (human)											
SOURCE	Homo sapiens											
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
REFERENCE	1 (bases 1 to 1228)											
AUTHORS	Rosen, C.A., Ruben, S.M., Ebner, R., Young, P.E., Ni, J., Moore, P.A., Komatsu, G. and Birse, C.E.											
TITLE	33 human secreted proteins											
JOURNAL	Patent: JP 2002534972-A 10 22-OCT-2002;											
COMMENT	HUMAN GENOME SCIENCES INC											
	OS Homo sapiens (human)											
	PN JP 2002534972-A/10											
	PD 22-OCT-2002											
	PF 18-JAN-2000 JP 2000594904											
	PR 19-JAN-1999 US 60/116330											
	PI CRAIG A ROSEN, STEVEN M RUBEN, REINHARD EBNER, PAUL E YOUNG, JIAN NI,											
	PI PAUL A MOORE, GEORGE KOMATSU, CHARLES E BIRSE PC											
	C12N15/09, A61K31/7115, A61K35/76, A61K38/00, A61K45/00, A61K48/00, PC											
	A61P1/00,											
	PC											
	A61P1/16, A61P3/10, A61P5/00, A61P7/02, A61P7/04, A61P7/06, A61P7/08, PC											
	A61P9/00,											
	PC A61P9/06, A61P9/08, A61P9/10, A61P9/12, A61P9/14, A61P11/00 PC											
	, A61P1/06, A61P13/02,											
	PC A61P13/12, A61P15/00, A61P17/02, A61P17/06, A61P19/02, A61P21/00,											
	PC A61P21/04,											
	PC A61P25/00, A61P25/14, A61P25/16, A61P25/28, A61P25/30, A61P27/02,											
	PC A61P27/06,											
	PC A61P29/00, A61P29/00, A61P31/04, A61P31/10, A61P31/12, A61P31/18,											
	PC A61P33/00,											
	PC A61P35/00, A61P35/02, A61P35/04, A61P37/02, A61P37/08, A61P39/02,											
	PC A61P43/00,											
	PC C07K34/47, C07K36/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC											
	C12P21/02, C12Q1/02,											
	PC C12Q1/68, G01N33/15, G01N33/50, G01N33/53,											

/mol_type="genomic DNA" /db_xref="taxon:9606"		99.9%; Score 1196; DB 6; Length 1228; 100.0%; Pred. No. 7.5e-286; Matches 1196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CAGCAGTGGTCTCTCAGTCCTCTCAAGCAGGAAAGAGTACTGTGTGCTGAGAGCCAT	60
Db	33	CAGCAGTGGTCTCTCAGTCCTCTCAAGCAGGAAAGAGTACTGTGTGCTGAGAGCCAT	92
QY	61	GGCAAGAATCCTCCAGAGAAATGTGAAGACTGTCAATTCTAAATCGAAGAGCTTTTAA	120
Db	93	GGCAAGAATCCTCCAGAGAAATGTGAAGACTGTCAATTCTAAATCGAAGAGCTTTTAA	152
QY	121	ATCCAAGAAATATGTAATCATTAAAGATTGTGGACTGGTGTGTTGGTATCTCTGGCCCT	180
Db	153	ATCCAAGAAATATGTAATCATTAAAGATTGTGGACTGGTGTGTTGGTATCTCTGGCCCT	212
QY	181	AATCTAATTGTCCTGTTTGGGGGCAAGCACTCTGGCCGGAGTACCCAAAAGC	240
Db	213	AATCTAATTGTCCTGTTTGGGGGCAAGCACTCTGGCCGGAGTACCCAAAAGC	272
QY	241	CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAAGATTTACATGGAAT	300
Db	273	CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAAGATTTACATGGAAT	332
QY	301	TGATCCTGTGACAGAACTGAATATTACAGACGGGAATGGCACTGATGAACATTTGGA	360
Db	333	TGATCCTGTGACAGAACTGAATATTACAGACGGGAATGGCACTGATGAACATTTGGA	392
QY	361	AGTGACAGACTTTAAAAACGGATACACTGCATCTACTTCGTGGGTCTTCAAAAATGTTT	420
Db	393	AGTGACAGACTTTAAAAACGGATACACTGCATCTACTTCGTGGGTCTTCAAAAATGTTT	452
QY	421	TATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACACAGAGAGAAATACATGA	480
Db	453	TATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACACAGAGAGAAATACATGA	512
QY	481	GAATGAAGAAATTTACACAACCTTTCTTGAACAGTCAAGTATTTGGTCCAGCAAAA	540
Db	513	GAATGAAGAAATTTACACAACCTTTCTTGAACAGTCAAGTATTTGGTCCAGCAAAA	572
QY	541	GCCTATTGAAAACCGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATACT	600
Db	573	GCCTATTGAAAACCGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATACT	632
QY	601	GACCATGTATTGGATCAATCCCACTCTAATATACAGTTTCTGAGTTTACAAGACTTTTGAGA	660
Db	633	GACCATGTATTGGATCAATCCCACTCTAATATACAGTTTCTGAGTTTACAAGACTTTTGAGA	692
QY	661	GGAGGAGAGATCTTCATCTTTCTGCGCAACGAAAAAAGGGATTGAACAAAATGAACA	720
Db	693	GGAGGAGAGATCTTCATCTTTCTGCGCAACGAAAAAAGGGATTGAACAAAATGAACA	752
QY	721	GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCGCTCACGCCACAGACAGTAGTAGGA	780
Db	753	GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCGCTCACGCCACAGACAGTAGTAGGA	812
QY	781	AGAACTTCCAATAAATGACTACTGAAATGGAATAGAAATTTGATCCATGCTCGATGA	840
Db	813	AGAACTTCCAATAAATGACTACTGAAATGGAATAGAAATTTGATCCATGCTCGATGA	872
QY	841	GAGAGGTTATTTGTAATTTACTGCCGTGAGGGCAACCGCTATTGCCCGCGCTGTGA	900
Db	873	GAGAGGTTATTTGTAATTTACTGCCGTGAGGGCAACCGCTATTGCCCGCGCTGTGA	932
QY	901	ACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTCG	960
Db	933	ACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTCG	992
QY	961	TGTCATCATGCTTCTAATCTGGTGGGTGGCCCGCATGCTGGGGGGGCTCTAATAGGCGCT	1020



REFERENCE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

QY	2	AGCAGTGGTCTCTCAGTCCCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCATG	61
Db			
QY	1	AGCAGTGGTCTCTCAGTCCCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCATG	60
Db			
QY	62	GCAAGAATCCTCCAGAGAAATGTGAAGACTGTGCATCTCTAAATGCAGAACTTTTAA	121
Db			
QY	61	GCAAGAATCCTCCAGAGAAATGTGAAGACTGTGCATCTCTAAATGCAGAACTTTTAA	120
Db			
QY	122	TCCAGAAATATGTAATCACTTAAGATTTGTGGACTGGTGTGGTATCTCGGCCCTA	181
Db			
QY	121	TCCAGAAATATGTAATCACTTAAGATTTGTGGACTGGTGTGGTATCTCGGCCCTA	180
Db			
QY	182	ACTCTAAATGTCTGTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAGCC	241
Db			
QY	181	ACTCTAAATGTCTGTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAGCC	240
Db			
QY	242	TATGACATGGAGACACTTCTACAGCAATGGAGAGAAAGATTTACATGGAAAT	301
Db			
QY	241	TATGACATGGAGACACTTCTACAGCAATGGAGAGAAAGATTTACATGGAAAT	300
Db			
QY	302	GATCCTGTGACAGAACTCAAAATATTCAGAAGCGAAATGGCACTGATCAAACTTGGAA	361
Db			
QY	301	GATCCTGTGACAGAACTCAAAATATTCAGAAGCGAAATGGCACTGATGAAACTTGGAA	360
Db			
QY	362	GTGACGACTTTTAAAAACGGATACACTGGCATCTACTTCGTGGTCTTCAAAAATGTTT	421
Db			
QY	361	GTGACGACTTTTAAAAACGGATACACTGGCATCTACTTCGTGGTCTTCAAAAATGTTT	420
Db			
QY	422	ATCAAACTCAGATTAAGTGATCTCGAATTTCTGAACCAAGAGAGAAATAGATGAG	481
Db			
QY	421	ATCAAACTCAGATTAAGTGATCTCGAATTTCTGAACCAAGAGAGAAATAGATGAG	480
Db			
QY	482	AATGAAGAAATACCAACAATTTCTTTGAACAGTCAAGTGAATTTGGTCCCGAGAAAAG	541
Db			
QY	481	AATGAAGAAATACCAACAATTTCTTTGAACAGTCAAGTGAATTTGGTCCCGAGAAAAG	540
Db			
QY	542	CCTATTGAAAACCGAGATTTCTTAAAAATTCAAAAATTTCTGGAGATTTGTGATAACGTG	601
Db			
QY	541	CCTATTGAAAACCGAGATTTCTTAAAAATTCAAAAATTTCTGGAGATTTGTGATAACGTG	600
Db			
QY	602	ACCATGTTTGGATCAATCCCACTCAATCAGTTTCTGAGTTACAAGACTTTTGGAG	661
Db			
QY	601	ACCATGTTTGGATCAATCCCACTCAATCAGTTTCTGAGTTACAAGACTTTTGGAG	660
Db			
QY	662	GAGGAGAGATCTTCACTTTCTGCTCAACGAAAAAAGGGATTTGAACAAAATGAACAG	721
Db			
QY	661	GAGGAGAGATCTTCACTTTCTGCTCAACGAAAAAAGGGATTTGAACAAAATGAACAG	720
Db			
QY	722	TGGGTGGTCCCTCAAGTGAAGTAGAGAACCCGTCACGCCAGAACAGAGTGAAGAA	781
Db			
QY	721	TGGGTGGTCCCTCAAGTGAAGTAGAGAACCCGTCACGCCAGAACAGAGTGAAGAA	780
Db			
QY	782	GAATTTCCAAATGAATGACTACTGAAATGGATAGAAATTTGATCCCATGCTGGATGAG	841
Db			
QY	781	GAATTTCCAAATGAATGACTACTGAAATGGATAGAAATTTGATCCCATGCTGGATGAG	840
Db			
QY	842	AGAGGTATTGTTGTTATTTACTCCGTCGAGGCAACCGCTATTTCGCGCCCGCTCTGTAA	901
Db			
QY	841	AGAGGTATTGTTGTTATTTACTCCGTCGAGGCAACCGCTATTTCGCGCCCGCTCTGTAA	900
Db			
QY	902	CTTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGGAGAGAGTCTGTCTGT	961
Db			
QY	901	CTTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGGAGAGAGTCTGTCTGT	960
Db			
QY	962	GTCAATCGCTTGTAACTGCTGGTGGCCCGCATGCTGGGAGGGTCTAATAGGAGTT	1021
Db			
QY	961	GTCAATCGCTTGTAACTGCTGGTGGCCCGCATGCTGGGAGGGTCTAATAGGAGTT	1020
Db			
QY	1022	TGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTTCAATGAATTT	1081
Db			
QY	1021	TGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTTCAATGAATTT	1080
Db			

QY	1082	CTGCTATGAGGCATCTGGCCCTGGTAGCCAGCTCTCAGAAATTAATCTTGTAGTAATTC	1141
Db			
QY	1081	CTGCTATGAGGCATCTGGCCCTGGTAGCCAGCTCTCAGAAATTAATCTTGTAGTAATTC	1140
Db			
QY	1142	CTCTCTTCATGTTCTAAATAAACTTCTACATATACCA	1179
Db			
QY	1141	CTCTCTTCATGTTCTAAATAAACTTCTACATATACCA	1178
Db			
RESULT 9			
AR338896			
LOCUS	AR338896	1380 bp	DNA
DEFINITION	Sequence 387 from patent US 6569662.		Linear
ACCESSION	AR338896		PAT 17-AUG-2003
VERSION	AR338896.1	GI:33725753	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1380)		
TITLE	Tang, Y.T., Zhou, P. and Drmanac, R.T.		
JOURNAL	Nucleic acids and polypeptides		
FEATURES	Patent: US 6569662-A 387 27-MAY-2003;		
	Location/Qualifiers		
	1..1380		
	source		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
	Query Match	98.1%;	Score 1174; DB 6; Length 1380;
	Best Local Similarity	99.6%;	Pred. No. 2.1e-280;
	Matches 1177; Conservative	0;	Mismatches 5; Indels 0; Gaps 0;
QY	1	CAGCAGTGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCAT	60
Db			
QY	73	CAGCAGGCGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCAT	132
Db			
QY	61	GGCAAGAATCCTCCAGAGAAATGTGAAGACTGTACATCTCTAAATGCAGAACTTTTAA	120
Db			
QY	133	GGCAAGAATGTGTCAGAGAAATGTGAAGACTGTACATCTCTAAATGCAGAACTTTTAA	192
Db			
QY	121	ATCCAGAAATATGTGAAATCACTTAAGATTTGTGAGCTGGTGTGGTATCTCGGCCCT	180
Db			
QY	193	ATCCAGAAATATGTGAAATCACTTAAGATTTGTGAGCTGGTGTGGTATCTCGGCCCT	252
Db			
QY	181	AACTCTAAATGTCTCTGTGGGGAGCAAGCACTTTCTGGCCGGAGGTACCCAAAAGC	240
Db			
QY	253	AACTCTAAATGTCTGTGGGGAGCAAGCACTTTCTGGCCGGAGGTACCCAAAAGC	312
Db			
QY	241	CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAGATTTACATGGAAT	300
Db			
QY	313	CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAGATTTACATGGAAT	372
Db			
QY	301	TGATCTGTGACCAAGCACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTTGA	360
Db			
QY	373	TGATCTGTGACCAAGCACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTTGA	432
Db			
QY	361	AGTGACGACTTTTAAAAAACGGATACACTGGCATCTACTCTGGTGGTCTTCAAAAATGTTT	420
Db			
QY	433	AGTGACGACTTTTAAAAAACGGATACACTGGCATCTACTCTGGTGGTCTTCAAAAATGTTT	492
Db			
QY	421	TATCAAACTCAGATTAAGTGATTTCTGAACTGATTTCTGACCAAGAGAGAAATAGATGA	480
Db			
QY	493	TATCAAACTCAGATTAAGTGATTTCTGAACTGATTTCTGAACTGATTTGAGTCCAGAGAA	552
Db			
QY	481	GAATGAAGAAATATACCAACTTTCTTTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGT	540
Db			
QY	553	GAATGAAGAAATATACCAACTTTCTTTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGT	612
Db			
QY	541	GCTATTGAAAACCGAGATTTCTTAAAAATTCAAAATTCGAGATTTGTGATAACGT	600
Db			
QY	613	GCTATTGAAAACCGAGATTTCTTAAAAATTCAAAATTCGAGATTTGTGATAACGT	672
Db			



61	GGCAAGAAATCTCTCCAGAGAAATGTGTGAAGACTGTGCACATTTCTAAATGCGACAGCGTTTAA	QY	860	TGTTGTATTTACTGCGGTGCGAGCGACCGCTATTTCGCCGGCGTCTGTGAACCTTTACTA	919
189	GGCAAGAAATCTCTCCAGAGAAATGTGTGAAGACTGTGCACATTTCTAAATGCGACAGCGTTTAA	Db	911	GGCTACTACCCATATCCATFACTGCTACCAAGAGAGACGAGTCATCTGTGTCATCATG	970
121	ATCCAGAAATATGTAAATCACTTAAAGATTTGTGGACTCGTGTGTGGTATCTCTGGCCCT	QY	920	GGCTACTACCCATATCCATFACTGCTACCAAGAGAGACGAGTCATCTGTGTCATCATG	979
249	ATCCAGAGAAATATGTAAATCACTTAAAGATTTGTGGACTCGTGTGTGGTATCTCTGGCCCT	Db	971	CTTTGTAATCTGGTGGCGCGCATGCTGGGAGGGTCTAATAGGAGTTTGAGCTCAA	1030
181	AACCTTAATTTGTCCTGTTTTGGGGGAGCAAGCACTTCTGCCCGGAGGTACCCAAAAAGC	QY	980	CTTTGTAATCTGGTGGCGCGCATGCTGGGAGGGTCTAATAGGAGTTTGAGCTCAA	1039
309	AACCTTAATTTGTCCTGTTTTGGGGGAGCAAGCACTTCTGCCCGGAGGTACCCAAAAAGC	Db	1031	ATGCTTAAATCTGCGCAACATATAATCAATCATCTATTCAATGAATTTCTGCGCTATG	1090
241	CTATGACATGAGACACACTTTTCTACAGCAATGGAGAGAGAGAGATTTACATGGAAAT	QY	1040	ATGCTTAAATCTGCGCAACATATAATCAATCATCTATTCAATGAATTTCTGCGCT-IG	1098
369	CTATGACATGAGACACACTTTTCTACAGCAATGGAGAGAGAGAGATTTACATGGAAAT	Db	1091	AGGCATCTGCCCTCGGTAGCCAGCTCTCCAGAAATTTAGTTAGGTAATTTCTCTCTTCA	1150
301	TGATCTGTGACAGAACTGAAATATTCTAGAAAGCGAAATGGCACTCATGAAACATTTGGA	QY	1099	AGGCATCTGCCCTCGGTAGCCAGCTCTCCAGAAATTTAGTTAGGTAATTTCTCTCTTCA	1158
429	TGATCTGTGACAGAACTGAAATATTCTAGAAAGCGAAATGGCACTCATGAAACATTTGGA	Db	1151	TGTTCTTAATAAATCTTACATATTATCACCACAAAAAATTTTTTTTTTTTTTTTTTTTT	1197
361	AGTGACGACATTTTAAAAACGGATACACTGGCATCTACTTCTGGGTCTTCAAAATGTTT	QY	1159	TGTTCTTAATAAATCTTACATATTATCACCACAAAAAATTTTTTTTTTTTTTTTTTTTT	1205
489	AGTACACACATTTTAAAAACGGATACACTGGCATCTACTTCTGGGTCTTCAAAATGTTT	Db			
421	TATCAAACTCAGATTTAAAGTGATCTCTGAATTTTCTGAACAGAGAGAGAAATAGATGA	QY			
549	TATCAAACTCAGATTTAAAGTGATCTCTGAATTTTCTGAACAGAGAGAGAAATAGATGA	Db			
481	GAATGAAGAAATTTACCAACATCTTCTTGAACAGTCACTGATTTGGGTCCCGACAGAAAA	QY			
609	GAATGAAGAAATTTACCAACATCTTCTTGAACAGTCACTGATTTGGGTCCCGACAGAAAA	Db			
541	GCTTATTCGAAACCGAGATTTTCTTAAAAATTCGAAATTCGAGATTTGTGATAAGCT	QY			
669	GCTTATTCGAAACCGAGATTTTCTTAAAAATTCGAAATTCGAGATTTGTGATAAGCT	Db			
601	GACCATGTATGGATCAATCCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTTGAGA	QY			
729	GACCATGTATGGATCAATCCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTTGAGA	Db			
661	GGAGGAGAGAGATCTTCACTTCTCCGCAACGAAAGGATTTGAACAAAAATGAACA	QY			
789	GGAGGAGAGAGATCTTCACTTCTCCGCAACGAAAGGATTTGAACAAAAATGAACA	Db			
721	GTGGGTGTCCTCAAGTGAAAGTAGAGAAAGCCCGTCAGCCAGACGAAGTGAAGA	QY			
849	GTGGGTGTCCTCAAGTGAAAGTAGAGAAAGCCCGTCAGCCAGACGAAGTGAAGA	Db			
781	AGAACTTCCAATAAATGACTATATCGAAATGGAATTTGATCCCATCTGGATGA	QY			
909	AGAACTTCCAATAAATGACTATATCGAAATGGAATTTGATCCCATCTGGATGA	Db			
841	GAGAGGTATTGTTGTTATTTACTGCGGTGAGGCAACCCCTATTGCGCGCGGTCTGTGA	QY			
969	GAGAGGTATTGTTGTTATTTACTGCGGTGAGGCAACCCCTATTGCGCGCGGTCTGTGA	Db			
901	ACCTTTTACTAGGCTACTTACCCATATCCATCTCTACCAAGGAGGACGAGTCATCTGTG	QY			
1029	ACCTTTTACTAGGCTACTTACCCATATCCATCTCTACCAAGGAGGACGAGTCATCTGTG	Db			
961	TGTCAATCATGCCCTTTGAACTGTTGGTGGCCCGCATGCTGGGGAGGGTCTAATAGAGCT	QY			
1089	TGTCAATCATGCCCTTTGAACTGTTGGTGGCCCGCATGCTGGGGAGGGTCTAATAGAGCT	Db			
1021	TTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAAATGATCATCTTCAATCAAT	QY			
1149	TTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAAATGATCATCTTCAATCAAT	Db			
1081	TCTGCTTATCAGGCACTCTGGCCCTCTGGTACCCAGCTCTCCAGAAATTTCTGTAGGTAAT	QY			
1209	TCTGCTTATCAGGCACTCTGGCCCTCTGGTACCCAGCTCTCCAGAAATTTCTGTAGGTAAT	Db			
1141	CCTCTCTTCATGTTCTAATAAATTTCTAATTAATGATCATCTTCAATCAAT	QY			

```
Db 1269 CCTCTCTTCAAGTCTTAATAAACTTCTACATTAACAATAA 1309
|||||
RESULT 12
AF291656 1184 bp mRNA linear PRI 07-DEC-2001
LOCUS Homo sapiens chondromodulin-IB mRNA, complete cds.
DEFINITION AF291656
ACCESSION AF291656
VERSION AF291656.1 GI:15077275
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1184)
Brandau, O., Meindl, A., Fassler, R. and Aszodi, A.
A novel gene, tendin, is strongly expressed in tendons and
ligaments and shows high homology with chondromodulin-I
Dev. Dyn. 221 (1), 72-80 (2001)
2125555
MEDLINE
PUBMED 11357195
REFERENCE 2 (bases 1 to 1184)
AUTHORS Brandau, O., Aszodi, A., Meindl, A. and Fassler, R.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2000) Experimental Pathology, Lund University,
Solvegatan 25, Lund 22185, Sweden
FEATURES
source Location/Qualifiers
1..1184
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
73..1026
/note="CHM-IB"
/codon_start=1
/product="chondromodulin-IB"
/protein_id="AAK83109.1"
/db_xref="GI:15077276"
/translators="MAKNPPENCEDCHLNABAFKSKIKCSLKI CGLVFGILALTLI
VLPFSKHFWPEPKYKAYMEHTFYSNGEKKIYMEIDPVTRIELFRSNGTDETLV
HDFKNGYTGIFVGLQKCFIKQIVPESEPEEIDNEEITTFEQSIVWVPAE
KPLENDFLKSLILEICDNVTWYNPTLISVSELDQFEEDGDLHFPANEKKGIEQ
NEQVVPQVKVETRHARQSEBELPINDYENGIEFDPLMDERGYCCIIYCRGRNYC
RRVCEPLLGYPPYCYQGGRVICRVIMPNNWVARMGRV"
ORIGIN
Query Match 97.78; Score 1170; DB 9; Length 1184;
Best Local Similarity 100.0%; Pred. No. 2.le-279;
Matches 1170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAGGAAGAGTACTGTGTGCTGAGAGCCAT 60
DB 15 CAGCAGTGGTCTCTCAGTCTCTCAAGCAGGAAGAGTACTGTGTGCTGAGAGCCAT 74
QY 61 GGCAAGAATCCTCCAGAGAATTGTGAAGACTGTACATTTCTAAATGCAGAAGCTTTAA 120
DB 75 GGCAAGAATCCTCCAGAGAATTGTGAAGACTGTACATTTCTAAATGCAGAAGCTTTAA 134
QY 121 ATCCAGAAATATGTAAATCATTAAAGATTGTGACTGTGTGGTATCTGGCCCT 180
DB 135 ATCCAGAAATATGTAAATCATTAAAGATTGTGACTGTGTGGTATCTGGCCCT 194
QY 181 AACTCTAATTGTCTCTGTTTGGGGGACAGACCTCTCTGGCCGAGGTACCCAAAAGC 240
DB 195 AACTCTAATTGTCTCTGTTTGGGGGACAGACCTCTCTGGCCGAGGTACCCAAAAGC 254
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTTACATGGAAAT 300
DB 255 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTTACATGGAAAT 314
QY 301 TGATCTGTGACCGAAGACTGAATATTTCAGAGCGGAAATGGCACTGATGAACATTGGA 360
DB 315 TGATCTGTGACCGAAGACTGAATATTTCAGAGCGGAAATGGCACTGATGAACATTGGA 374
```

```
QY 361 AGTGACGAGCTTTAAAAACGATACACTGGCATCTACTCTGTGGGTCTTCAAAATGTTT 420
DB 375 AGTGACGAGCTTTAAAAACGATACACTGGCATCTACTCTGTGGGTCTTCAAAATGTTT 434
QY 421 TATCAAAACTCAGATTAAAGTGAATTCCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 480
DB 435 TATCAAAACTCAGATTAAAGTGAATTCCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 494
QY 481 GAATGAAGAAATTACCAAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCCAAGAGAAA 540
DB 495 GAATGAAGAAATTACCAAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCCAAGAGAAA 554
QY 541 GCCTATTGAAAAACCGAGATTTCTTAAAAATTCCAAATTTCTGGAGATTTGTGATAACGT 600
DB 555 GCCTATTGAAAAACCGAGATTTCTTAAAAATTCCAAATTTCTGGAGATTTGTGATAACGT 614
QY 601 GACCATGTATTGGATCAATCCCACTCTTAATATCAGTTTTCTGAGTTACAAGACTTTGAGGA 660
DB 615 GACCATGTATTGGATCAATCCCACTCTTAATATCAGTTTTCTGAGTTACAAGACTTTGAGGA 674
QY 661 GGAGGGAGAAAGATCTTCACCTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACA 720
DB 675 GGAGGGAGAAAGATCTTCACCTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACA 734
QY 721 GTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGCAAGTGAAGGA 780
DB 735 GTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGCAAGTGAAGGA 794
QY 781 AGAACTTCCAAATAATGACTATATCTGAAAAATGAATAAATTTGATCCCATGCTGGAATGA 840
DB 795 AGAACTTCCAAATAATGACTATATCTGAAAAATGAATAAATTTGATCCCATGCTGGAATGA 854
QY 841 GAGAGGTTATTGTGTTATTTACTGCGCTCGAGGCAACCGCTATTGCGCGCGCGTCTGTGA 900
DB 855 GAGAGGTTATTGTGTTATTTACTGCGCTCGAGGCAACCGCTATTGCGCGCGCGTCTGTGA 914
QY 901 ACCTTTACTAGGCTACTACCCATATCCATCTCTACCAAGAGGAGGAGTCAATCTGTGCG 960
DB 915 ACCTTTACTAGGCTACTACCCATATCCATCTCTACCAAGAGGAGGAGTCAATCTGTGCG 974
QY 961 TGTCAATCATGCCCTTTGTAACCTGTGGGTGGCGCGCATGCTCGGGAGGGTCTTAATAGAGGT 1020
DB 975 TGTCAATCATGCCCTTTGTAACCTGTGGGTGGCGCGCATGCTCGGGAGGGTCTTAATAGAGGT 1034
QY 1021 TTGAGCTCAAAATGCTTTAACTGCTGGCAACATATATAAATGATGCTATTCATGAATTT 1080
DB 1035 TTGAGCTCAAAATGCTTTAACTGCTGGCAACATATATAAATGATGCTATTCATGAATTT 1094
QY 1081 TCTGCTTATGAGGCACTTGCGCCCTGTGTAGCCAGTCTCCAGAAATTAATTTGTAGGTAAT 1140
DB 1095 TCTGCTTATGAGGCACTTGCGCCCTGTGTAGCCAGTCTCCAGAAATTAATTTGTAGGTAAT 1154
QY 1141 CCTCTCTTCATGTTCTAAATAAACTTCTACA 1170
DB 1155 CCTCTCTTCATGTTCTAAATAAACTTCTACA 1184
RESULT 13
AB055421 1140 bp mRNA linear PRI 07-FEB-2001
LOCUS Homo sapiens mRNA for CHML, complete cds.
DEFINITION AB055421
ACCESSION AB055421
VERSION AB055421.1 GI:12698292
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Yamana, K., Wada, H., Takahashi, Y., Sato, H., Kasahara, Y. and
Kiyoki, M.
Molecular cloning and characterization of CHML, a novel membrane
```



molecule similar to chondromodulin-I  
Biochem. Biophys. Res. Commun. 280 (4), 1101-1106 (2001)  
21092728  
11162640  
2 (bases 1 to 1140)  
Direct Submission  
Submitted (05-FEB-2001) Kei Yamana, Teijin Limited, Teijin  
Institute for Biomedical Research; 4-3-2 Asahigaoka, Hino, Tokyo  
191-8512, Japan (E-mail:k.yamana@teijin.co.jp,  
Tel:81-42-586-8281(ex.8281), Fax:81-42-587-5519)  
Tel:81-42-586-8281(ex.8281), Fax:81-42-587-5519)  
FEATURES  
source  
1..1140  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..1140  
/gene="CHM1L"  
7..960  
/gene="CHM1L"  
/codon\_start=1  
/product="ChM1L"  
/protein\_id="BAB21756.1"  
/db\_xref="GI:12698293"  
/translation="WAKNPENCEDCHIINAEAFKSKTKSLKICGLVFLALTLI  
VLFWSGKHPWEPVKKAYDMHTFYSGKEPKIYMEIDVETREIFRSNGTDETLV  
HDFKNGYTGIVFGLQKCFIKQIKVIFPESPEBEIDENEETITTFPQSVIWPAAE  
KPIENRDLKSLILEICDNVTWYNPTLISVSELODPEEEDLHPFANEKKGLEQ  
NQRWVPOVKVETKHARQASBEELPINDYTENGIEFDPLDERGYCCYICRRGNRYC  
RVCPELLGYFPYCYQGRVICRVIMPCNWWVARMGRV"  
ORIGIN  
Query Match 95.1%; Score 1138.4; DB 9; Length 1140;  
Best Local Similarity 99.9%; Pred. No. 1.5e-271;  
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 53 GAGACATGCGAAGAAATCTCCAGAGAATTGTGAAGACTGTGCACATTCTAATGCGAGAA 112  
DB 1 GAGACATGCGAAGAAATCTCCAGAGAATTGTGAAGACTGTGCACATTCTAATGCGAGAA 60  
QY 113 GCTTTTAAATCCAGAAATATGTAATCACTTAAGATTGTGACTGTGTGTGTATC 172  
DB 61 GCTTTTAAATCCAGAAATATGTAATCACTTAAGATTGTGACTGTGTGTGTATC 120  
QY 173 CTGGCCCTAACTCTAATCTCCTGTTTGGGGAGCAAGCACTCTGCGCGAGGTACCC 232  
DB 121 CTGGCCCTAACTCTAATCTCCTGTTTGGGGAGCAAGCACTCTGCGCGAGGTACCC 180  
QY 233 AAAAAGCCTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAAAGATTAC 292  
DB 181 AAAAAGCCTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAAAGATTAC 240  
QY 293 ATGGAATGATCCTGTGACCAAGCACTGAAATATCAGAAAGCGAAATGGCACTGATGAA 352  
DB 241 ATGGAATGATCCTGTGACCAAGCACTGAAATATCAGAAAGCGAAATGGCACTGATGAA 300  
QY 353 ACATTGGAAGTGCAGACTTTTAAAAACCGATACATGGCATCTACTTCGTGGGTCTTCAA 412  
DB 301 ACATTGGAAGTGCAGACTTTTAAAAACCGATACATGGCATCTACTTCGTGGGTCTTCAA 360  
QY 413 AAATGTTTTATCAAACTCAGATTAAAGTGATCTCGAATTTCTGAACAGAGAGAA 472  
DB 361 AAATGTTTTATCAAACTCAGATTAAAGTGATCTCGAATTTCTGAACAGAGAGAA 420  
QY 473 ATAGATGAGANTGAAGAAATATACCAAACTTTCTTTGAAACAGTCACTGATTTGGGTCCCA 532  
DB 421 ATAGATGAGANTGAAGAAATATACCAAACTTTCTTTGAAACAGTCACTGATTTGGGTCCCA 480  
QY 533 GCAGAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCGAGAGATTGT 592  
DB 481 GCAGAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCGAGAGATTGT 540  
QY 593 GATAACGTGACCATGATTGATGATCAATCCCACTCTAATATACATTTCTGAGTTTACAAGAC 652

Db 541 GATAACGTGACCATGATTGATCAATCCCACTCTAATATCACTTTCTGAGTTACAAGAC 600  
QY 653 TTTTGGAGGAGGAGAGAGATCTTCACTTTTCTGCGCAACGAAAAAAGGATTTGAACAA 712  
DB 601 TTTTGGAGGAGGAGAGAGATCTTCACTTTTCTGCGCAACGAAAAAAGGATTTGAACAA 660  
QY 713 AATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCACGCCAGACAGCA 772  
DB 661 AATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCACGCCAGACAGCA 720  
QY 773 AGTGAGGAAGAACTTCCAAATGAATGACTATGTAATGGAATGGAATTTGATCCCATG 832  
DB 721 AGTGAGGAAGAACTTCCAAATGAATGACTATGTAATGGAATGGAATTTGATCCCATG 780  
QY 833 CTGGATGAGAGAGGTTATTGTTGTTTACTGCGCTCGAGGCAACCGCTATTGTCGCCCGC 892  
DB 781 CTGGATGAGAGAGGTTATTGTTGTTTACTGCGCTCGAGGCAACCGCTATTGTCGCCCGC 840  
QY 893 GTCTGTGAACCTTTACTAGGCTACTACCCATATCCATATGCTACCAAGAGGACGAGTC 952  
DB 841 GTCTGTGAACCTTTACTAGGCTACTACCCATATCCATATGCTACCAAGAGGACGAGTC 900  
QY 953 ATCTGTCGTGTCATCATGCTTGTAACTGCTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012  
DB 901 ATCTGTCGTGTCATCATGCTTGTAACTGCTGGTGGCGCGCATGCTGGGAGGGTCTAA 960  
QY 1013 TAGGAGGTTTGAAGCTCAAAATGCTTAAACTGCTGGCAACATAATAAATGATGCTATTTC 1072  
DB 961 TAGGAGGTTTGAAGCTCAAAATGCTTAAACTGCTGGCAACATAATAAATGATGCTATTTC 1020  
QY 1073 AATGAATTTCTGCTATGAGGCACTGGCCCTGCTAGGACGCTCTCCAGAAATTACTGT 1132  
DB 1021 AATGAATTTCTGCTATGAGGCACTGGCCCTGCTAGGACGCTCTCCAGAAATTACTGT 1080  
QY 1133 AGTAAATTCCTCTCTTCATGTTCTAATAAATCTTACATTATCACCACAAAAAATAA 1192  
DB 1081 AGTAAATTCCTCTCTTCATGTTCTAATAAATCTTACATTATCACCACAAAAAATAA 1140  
RESULT 14  
AR338895  
LOCUS AR338895 1428 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 386 from patent US 6569662.  
ACCESSION AR338895  
VERSION AR338895.1 GI:33725752  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1428)  
AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.  
TITLE Nucleic acids and polypeptides  
JOURNAL Patent: US 6569662-A 386 27-MAY-2003;  
FEATURES Location/Qualifiers  
source  
1..1428  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 93.2%; Score 1116; DB 6; Length 1428;  
Best Local Similarity 95.7%; Pred. No. 5.3e-266;  
Matches 1177; Conservative 0; Mismatches 5; Indels 48; Gaps 1;  
QY 1 CAGCAGTGGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTGTGAGAGACCAT 60  
DB 73 CAGCAGGCTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTGTGAGAGACCAT 132  
QY 61 GCGAAAGATCTCTCAGAGAAATGGAAGACTGTCACTTCTAAATGCAGAGCTTTTAA 120  
DB 133 GCGAAAGATGCTCCAGAGAAATGGAAGACTGTCACTTCTAAATGCAGAGCTTTTAA 192  
QY 121 ATCCAGAAATATGTAATGTAATCACTTAAGATTTGTGAGCTGGTGTGTTGGTATCCTGGCCCT 180



Db 193 ATCCAAGAAATATGTAAATCATTTAAGATTGTGGACTGGTGTGGTATCTCTGGCCCT 252  
QY 181 AACTCTAATGTCCTGTTTGGGGAGCAGACATTTCTGGCCGGAGGTACCCAAAAAGC 240  
Db 253 AACTCTAATGTCCTGTTTGGGGAGCAGACATTTCTGGCCGGAGGTACCCAAAAAGC 312  
QY 241 CTATGACATGAGCACACTTTCTACAGCAATGGAGAGAGAGAGATTTCATCATGAAT 300  
Db 313 CTATGACATGAGCACACTTTCTACAGCAATGGAGAGAGAGAGATTTCATCATGAAT 372  
QY 301 TGATCTGTGACCGAATCTGAATATTCAGAGCGGAATGGCACTGATGAACATTGGA 360  
Db 373 TGATCTGTGACCGAATCTGAATATTCAGAGCGGAATGGCACTGATGAACATTGGA 432  
QY 361 AGTCACGACCTTAAACACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAATGTTT 420  
Db 433 AGTACACGACCTTAAACACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAATGTTT 492  
QY 421 TATCAAAACTCAGATTAAAGTGATTCCTGAATTTTCTGAACCCAGAGAGAAATAGATGA 480  
Db 493 TATCAAAACTCAGATTAAAGTGATTCCTGAATTTTCTGAACCCAGAGAGAAATAGATGA 552  
QY 481 GAATGAAGAAATACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCGACGAGAAA 540  
Db 553 GAATGAAGAAATACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCGACGAGAAA 612  
QY 541 GCCTATTGAACACCGAGATTTCCTTAAAAATTCAGAAATCTGGAGATTGTGTAAACGT 600  
Db 613 GCCTATTGAACACCGAGATTTCCTTAAAAATTCAGAAATCTGGAGATTGTGTAAACGT 672  
QY 601 GACCATGTATTGGATCAATCCCACTCTAATATC----- 633  
Db 673 GACCATGTATTGGATCAATCCCACTCTAATATCAGAAACATTTGCAAGCAGTTGCATCA 732  
QY 634 -----AGTTTCTGAGTTTCAAGACTTTTGAGGAGGGGAGAGA 672  
Db 733 CAACCTTTGCAATTTATTCTTAGTTTCTGAGTTTCAAGACTTTTGAGGAGGGGAGAGA 792  
QY 673 TCTTCACCTTCTGCGACGAGAAAGAGGATTTGAACAAATGACAGTGGTGGTCC 732  
Db 793 TCTTCACCTTCTGCGACGAGAAAGAGGATTTGAACAAATGACAGTGGTGGTCC 852  
QY 733 TCAAGTGAAGTAGAGAAGACCCGTCAGCCAGACCAAGTAGTAGGAGAACTTCCAAAT 792  
Db 853 TCAAGTGAAGTAGAGAAGACCCGTCAGCCAGACCAAGTAGTAGGAGAACTTCCAAAT 912  
QY 793 AAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTATTG 852  
Db 913 AAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTATTG 972  
QY 853 TTGTATTACTGCGGTGAGGCAACCGCTATTGCGCGCGCGTCTGTGAACCTTTACTAGG 912  
Db 973 TTGTATTACTGCGGTGAGGCAACCGCTATTGCGCGCGCGTCTGTGAACCTTTACTAGG 1032  
QY 913 CTACTACCATATCATACTGCTACCAAGGAGGAGAGTCACTCTGTCGTGTCATCATGCC 972  
Db 1033 CTACTACCATATCATACTGCTACCAAGGAGGAGAGTCACTCTGTCGTGTCATCATGCC 1092  
QY 973 TTGTAACGTGGTGGTGGCCCGATGCTGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAT 1032  
Db 1093 TTGTAACGTGGTGGTGGCCCGATGCTGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAT 1152  
QY 1033 GCTTAAACTGTGGCAACATATAAATGATGCTGTTTCAATGAATTTCTGCCTATGAG 1092  
Db 1153 GCTTAAACTGTGGCAACATATAAATGATGCTGTTTCAATGAATTTCTGCCTATGAG 1212  
QY 1093 GCATCTGCGCCCTGGTAGCAGCTCTCCAGAAATTAATTTGTAGGTAAATTCCTCTTCATG 1152  
Db 1213 GCATCTGCGCCCTGGTAGCAGCTCTCCAGAAATTAATTTGTAGGTAAATTCCTCTTCATG 1272  
QY 1153 TTCTAATAAACTTCTACATTATACCAAAA 1182

1273 TTCTAATAAACTTCTACATTATCACAACA 1302

RESULT 15  
AB059407  
LOCUS 1175 bp mRNA linear MAM 19-SEP-2002  
DEFINITION Equus caballus TNMD mRNA for tenomodulin, complete cds.  
ACCESSION AB059407  
VERSION AB059407.2 GI:23200576  
KEYWORDS  
SOURCE  
ORGANISM  
EQUUS caballus (horse)  
EQUUS caballus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Perissodactylia; Equidae; Equus.  
REFERENCE  
1 Hasegawa,T.  
Molecular cloning and characterization of mRNA for equine  
tenomodulin  
Published Only in Database (2001)  
2 (bases 1 to 1175)  
Hasegawa,T.  
Direct Submission  
Submitted (05-APR-2001) Telhisa Hasegawa, JRA Equine Research  
Institute, Laboratory of Molecular and Cellular Biology; 321-4  
Tokami-Cho, Utsunomiya, Tochigi 320-0856, Japan  
(E-mail:telhisa@center.equinist.go.jp, Tel:81-28-647-0662,  
Fax:81-28-647-0686)  
COMMENT  
On Sep 19, 2002 this sequence version replaced gi:15982572.  
FEATURES  
source  
1..1175  
/organism="Equus caballus"  
/mol\_type="mRNA"  
/strain="thoroughbred"  
/db\_xref="taxon:9796"  
/tissue\_type="tendon"  
/dev\_stage="adult"  
/note="domestic horse"  
1..1175  
/gene="TNMD"  
52..1005  
/gene="TNMD"  
/codon\_start=1  
/product="tenomodulin"  
/protein\_id="BAB69482.2"  
/db\_xref="GI:23200577"  
/translation="MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTLI  
VLPFGKHFPETPKTYDMEHTYSNGEKKIYMEIDPVTRTIFRSGNQTDETLV  
HDFKNGYTGIVFVGLQKCFIKTKIVTPEFSEPEEEDNEETITTFEQGVIVWPAE  
KPIENRPLKNSKILEICDNVTWYNPTLIAVSELODFREDGDLHPTNDKKGIEQ  
NEQWVPOVKVETRRARQASEELPINDYTENGIEFDPMLDERGYCCICYCRGNRYC  
RRVEPLILGYPPYCYQGGRVICRVIMPCNWNWVARMGLRV"  
polyA\_signal  
1154..1159  
/gene="TNMD"  
1175  
polyA\_site  
/gene="TNMD"  
ORIGIN  
Query Match 87.2%; Score 1043.6; DB 4; Length 1175;  
Best Local Similarity 93.8%; Pred. No. 4.9e-248;  
Matches 1099; Conservative 0; Mismatches 69; Indels 4; Gaps 1;  
QY 8 GGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACCATGCCAAG 67  
Db 1 GGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACCATGCCAAG 60  
QY 68 AATCCTCCAGAGAAATGTGAAGACTGTCAATTCATAATGCAGAGCTTTTAAATCCAAG 127  
Db 61 AATCCTCCAGAGAACTGTGAGGACTGTCAATTTAAATGCAGAGCTTTTAAATCCAAG 120  
QY 128 AAAATATGTAATCACTTAAAGATTGTGGACTGTGTGGTATCTCTGGCCCTAACTCTA 187  
Db 121 AAGATATGTAATCACTTAAATTTGTGGATTGTGTGGTATCTCTCTGGCCCTAACTCTA 180

QY 188 ATTGTCCTCTTTGGGGAGCAAGCACTTCTGCGGAGGTACCCAAAAAGCCTATGAC 247  
Db |||||  
181 ATTGTCCTCTTTGGGGAGCAAGCACTTCTGCGGAGGTACCCAAAAAGCCTATGAC 240  
QY 248 ATGGAGCACACTTTTACAGCATGAGAGAGAAAGATTTTACATGGAATTTGATCCT 307  
Db |||||  
241 ATGGAGCACACTTTTACAGCAACGGAGAGAGAGAAATTTTACATGGAATTTGATCCC 300  
QY 308 GTGACCAGAACTGAAATATTTCAGAAGCGGAAATGGCACTGATGAAACATTTGGAAGTGCAC 367  
Db |||||  
301 GTGACCAGAACTGAAATATTTCAGAAGCGGAAATGGCACTGATGAAACATTTGGAAGTGCAC 360  
QY 368 GACTTTAAAAACCGATACACTGCACTCTACTCTGCGGCTTCAAAAATGTTTTATCAAA 427  
Db |||||  
361 GACTTTAAAAATGGATATACTGCACTCTACTCTGCGGCTTCAAAAATGTTTTATCAAA 420  
QY 428 ACTCAGATTAAAGTGATTTCTGATTTTCTGAACCGAGAGGAATAGATGAATGAA 487  
Db |||||  
421 ACTCAGATTAAAGTGATTTCTGATTTTCTGAACCGAGAGGAATAGATGAATGAA 480  
QY 488 GAAATTACCACTTTCTTTGAACAGTCACTGATTTGGGTCCAGCAGAAAAAGCCTATT 547  
Db |||||  
481 GAAATTACCACTTTCTTTGAACATCACTGATTTGGGTCCAGCAGAAAAAGCCTATT 540  
QY 548 GAAAAACCGAGATTTCTTAAAAATTCAAAAATTTCTGAGATTTGTGATAACGTGACCATG 607  
Db |||||  
541 GAAAAACCGAGATTTCTTAAAAATTCAAAAATTTCTGAGATTTGTGATAACGTGACCATG 600  
QY 608 TATTGGATCAATCCCACTCTAAATCACTTCTGAGTTTCAAGACTTTGAGGAGGAGGA 667  
Db |||||  
601 TATTGGATCAATCCCACTCTAAATCACTTCTGAGTTTCAAGACTTTGAGGAGGAGGT 660  
QY 668 GAAGATCTTCACTTTCTGCAACGAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 727  
Db |||||  
661 GAAGATCTTCACTTTCTGCAACGAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 720  
QY 728 GTCCTCTCAAGTGAAGTAGAGAGACCCGTCAGCCAGACAGCAAGTGAAGAGAACTT 787  
Db |||||  
721 GTCCTCTCAAGTGAAGTGAAGTAGAGACCCGTCAGCCAGACAGCAAGTGAAGAGAACTT 780  
QY 788 CCAATAATGACTATCTGAAATGGAATAGATTTGATCCCATGCTGGATGAGAGAGT 847  
Db |||||  
781 CCAATTAATGACTATCTGAAATGGAATAGATTTGATCCCATGCTGGATGAGAGAGT 840  
QY 848 TATTGTTGATTTACTGCCGTCGAGGCAACCGCTATTTCGCGCGCGTCTGTGAACCTTTA 907  
Db |||||  
841 TATTGTTGATTTACTGCCGTCGAGGCAACCGCTATTTCGCGCGCGTCTGTGAACCTTTA 900  
QY 908 CTAGGCTACTACCCATATCCATCTGCTACAGGAGGACGATCATCTCTGTCGTGATC 967  
Db |||||  
901 CTAGGTTACTACCCGTTATCCATCTGCTACAGGAGGCGGGTTATCTCTGTCGTGATC 960  
QY 968 ATGCTTTGTAACGTGGTGGCGCGCATGCTGGGAGGGTCTAATAGGAGGTTTGAGCT 1027  
Db |||||  
961 ATGCTTTGCAACTGGTGGTGGCGCGCATGTTGGGAGGGTCTAGTAGCAAGTTCGAGCA 1020  
QY 1028 CAAATGCTTAACTGCTG---CAACATATAATAATGCAATGCTATTTCAATGAATTTCT 1083  
Db |||||  
1021 CAAGAGCTTAAATTTCTGCGAGCAACATATAATAATGCAATGCTATTTCAATGAATTTCT 1080  
QY 1084 GCCTATGAGGATCTGCGCCCTGGTAGCCAGCTCTCAGAAATTAATTTGTTAGGTAATTCCT 1143  
Db |||||  
1081 CCCTATAAGGCAATTTGGCTCTGTTAGCCAGTACTCCAGAAATTACTTTGTAGGTAATTCCT 1140  
QY 1144 CTCTTCATGTTCTAATAAATTTCTACATATC 1175  
Db |||||  
1141 CTCCTCATGTTCTAATAAATTTCTACATATC 1172